


```

Db      1308  AATTGGCTCATTCATTCATGAGCCACTGTCGACTTTCTTCGAGGAGTCTGATCAGCC 1367
Qy      241  TACAGTATGGTGAATGCTGAGCAGCTGTGTGATTTTGGAGAAAGCTCTCTGTAAT 300
Db      1368  TACAGTATGGTGAATGCTGAGCAGCTGTGTGATTTTGGAGAAAGCTCTCTGTAAT 1427
Qy      301  CACACAAGCAGCAGCATTTATGCTGAGTCAAGCTCCATTTTCATTGTGCTTCATCTCC 360
Db      1428  CACACAAGCAGCAGCATTTATGCTGAGTCAAGCTCCATTTTCATTGTGCTTCATCTCC 1487
Qy      361  ATTGACCGCTACATATGCTGTGTGATCCACTGAGATATTAAGCCAAAGATATCTTG 420
Db      1488  ATTGACCGCTACATATGCTGTGTGATCCACTGAGATATTAAGCCAAAGATATCTTG 1547
Qy      421  GTTATTTGCTGATGATCTTCATTAGTTGGAGTGCCCGCTGTTTTCATTGGAATG 480
Db      1548  GTTATTTGCTGATGATCTTCATTAGTTGGAGTGCCCGCTGTTTTCATTGGAATG 1607
Qy      481  ATCTTTCTGAGAGCTAACTTCAAAAGCGCTGAAGAGATATATTACAAAGTTCATCTC 540
Db      1608  ATCTTTCTGAGAGCTAACTTCAAAAGCGCTGAAGAGATATATTACAAAGTTCATCTC 1667
Qy      541  AAGAGAGGTGCTGCTCTCTCTTTTACCAAAATATCTGGGCTACTGACCTTTATGACTCT 600
Db      1668  AAGAGAGGTGCTGCTCTCTCTTTTACCAAAATATCTGGGCTACTGACCTTTATGACTCT 1727
Qy      601  TTTTATATACCTGATCTATTTATGTTATGTCTATTTACAGAAATATCTTATGCTTAA 660
Db      1728  TTTTATATACCTGATCTATTTATGTTATGTCTATTTACAGAAATATCTTATGCTTAA 1787
Qy      661  GAACGAGGCAAGTATATAGTATGTCGCAATCAGAAAGCTCCAAATTGGATTGGAAATGAA 720
Db      1788  GAACGAGGCAAGTATATAGTATGTCGCAATCAGAAAGCTCCAAATTGGATTGGAAATGAA 1847
Qy      721  AATGGAATTTCAAAAGCAAGAAAGAAAGCTGGAAGACATTTGGGAGTGTGTGGA 780
Db      1848  AATGGAATTTCAAAAGCAAGAAAGAAAGCTGGAAGACATTTGGGAGTGTGTGGA 1907
Qy      781  GTTTTCTTAATATGCTGTGCTGCTTTCTTATCTGTACAGTCATGAGACCTTTTCTTAC 840
Db      1908  GTTTTCTTAATATGCTGTGCTGCTTTCTTATCTGTACAGTCATGAGACCTTTTCTTAC 1967
Qy      841  TACATATTTCCACCACTTATGATGATGTTGATTTGGTGTACTGTAAGTCTTACA 900
Db      1968  TACATATTTCCACCACTTATGATGATGATTTGGTGTACTGTAAGTCTTACA 2027
Qy      901  TTTAATCCAAATGTTATGATTTTCTATCTTGTGTTAGAAAGCACTGAAGATGATG 960
Db      2028  TTTAATCCAAATGTTATGATTTTCTATCTTGTGTTAGAAAGCACTGAAGATGATG 2087
Qy      961  CTGTTTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAATATTTTGGAAATGAGT 1020
Db      2088  CTGTTTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAATATTTTGGAAATGAGT 2147
Qy      1021  TCATAGAAATTTATATT 1038
Db      2148  TCATAGAAATTTATATT 2165

```

```

RESULT 2
AL513524/c 141028 bp DNA linear PRI 11-MAY-2001
LOCUS      Human DNA sequence from clone RP11-295F4 on chromosome 6, complete
DEFINITION
ACCESSION  AL513524
VERSION     AL513524.8  GI:14031112
KEYWORDS   HTG.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 141028)

```

AUTHORS TITLE JOURNAL COMMENT

Clark, G.
Direct Submission
Submitted (10-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On May 14, 2001 this sequence version replaced gi:13751552.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPep; Information on the WormPep
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chromosome6.html
RP11-295F4 is from the library RPCT-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/Dacpac/home.htm
VECTOR: DBACe3.6

FEATURES

source

1..141028
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-295F4"
/clone_1b="RPCT-11.2"
52..267
/note="L2 repeat: matches 1314..1547 of consensus"
97..1029
/note="L2 repeat: matches 2669..2701 of consensus"
1493..1562
/note="2 copies 35 mer 95% conserved"
2086..2215
/note="MER94 repeat: matches 1..133 of consensus"
2441..2610
/note="L1MD3 repeat: matches 7578..7738 of consensus"
2616..2704
/note="L1MD3 repeat: matches 6701..6788 of consensus"
2717..2832
/note="AluYb repeat: matches 3..124 of consensus"
2833..3295
/note="AluYb repeat: matches 1..444 of consensus"
3296..3497
/note="AluYb repeat: matches 124..304 of consensus"
3809..3844
/note="L1MC/D repeat: matches 5568..5603 of consensus"
3881..4045
/note="L1MC4 repeat: matches 6460..6611 of consensus"
4046..4111
/note="MADE1 repeat: matches 1..80 of consensus"
4112..4885
/note="L1MC4 repeat: matches 5600..6460 of consensus"
5473..5887

/note="LIMC4 repeat: matches 5201. .5630 of consensus"
5888. .6069
/note="PAM repeat: matches 1. .173 of consensus"
repeat_region 6070. .7026
/note="LIMC4 repeat: matches 4226. .5201 of consensus"
repeat_region 7454. .9534
/note="LIMC4 repeat: matches 2110. .4239 of consensus"
repeat_region 9835. .9963
/note="LIMC4 repeat: matches 1968. .2101 of consensus"
repeat_region 10234. .10331
/note="LTR16A repeat: matches 59. .153 of consensus"
repeat_region 10731. .11121
/note="MER57B repeat: matches 1. .403 of consensus"
repeat_region 11122. .15589
/note="MER57-internal repeat: matches 1. .4612 of consensus"
repeat_region 15590. .15899
/note="AluY repeat: matches 1. .306 of consensus"
repeat_region 15900. .16478
/note="MER57-internal repeat: matches 4612. .5175 of consensus"
repeat_region 16480. .16587
/note="AluSg/x repeat: matches 200. .307 of consensus"
repeat_region 17166. .17532
/note="THE1B repeat: matches 1. .364 of consensus"
repeat_region 17728. .17925
/note="L2 repeat: matches 1636. .1836 of consensus"
repeat_region 18077. .18479
/note="THE1B-INTERNAL repeat: matches 1186. .1574 of consensus"
repeat_region 18485. .18845
/note="THE1B repeat: matches 1. .364 of consensus"
repeat_region 18846. .18885
/note="L1PA15 repeat: matches 6115. .6157 of consensus"
repeat_region 19862. .19979
/note="LIMC4 repeat: matches 7857. .7977 of consensus"
repeat_region 20016. .21137
/note="MER11A repeat: matches 1. .1266 of consensus"
repeat_region 21138. .21336
/note="LIMC4 repeat: matches 7654. .7844 of consensus"
repeat_region 21337. .21608
/note="AluY repeat: matches 39. .293 of consensus"
repeat_region 21609. .22560
/note="LIMC4 repeat: matches 6650. .7654 of consensus"
repeat_region 23089. .23122
/note="L2 copies 2 mer ac 82% conserved"
repeat_region 28852. .28920
/note="MIR repeat: matches 8. .77 of consensus"
repeat_region 29765. .29800
/note="L18 copies 2 mer ca 97% conserved"
repeat_region 30361. .30361
/note="Sequence from overlapping clone ba80A23 (AL589691). Assembly confirmed by restriction digest."
misc_feature 30587. .30722
/note="HERV1 repeat: matches 3314. .3453 of consensus"
repeat_region 31082. .31233
/note="HERV1 repeat: matches 2705. .2858 of consensus"
repeat_region 31761. .31969
/note="MLT1H repeat: matches 29. .294 of consensus"
repeat_region 32170. .32723
/note="MER67C repeat: matches 1. .506 of consensus"
repeat_region 33639. .33922
/note="AluX repeat: matches 3. .286 of consensus"
repeat_region 34841. .35352
/note="MLT1D repeat: matches 1. .505 of consensus"
misc_feature 35113. .35279
/note="Sequence from overlapping clone ba80A23 (AL589691). Assembly confirmed by restriction digest"
repeat_region 35703. .36065
/note="L2 repeat: matches 1141. .1515 of consensus"
repeat_region 36191. .36282
/note="46 copies 2 mer tt 62% conserved"
repeat_region 36321. .36350

/note="L5 copies 2 mer tt 86% conserved"
repeat_region 36450. .36652
/note="L2 repeat: matches 1675. .1914 of consensus"
repeat_region 37308. .38171
/note="MER45C repeat: matches 1. .895 of consensus"
repeat_region 40257. .40551
/note="AluSg repeat: matches 20. .313 of consensus"
repeat_region 40619. .40973
/note="MLT1B repeat: matches 1. .390 of consensus"
repeat_region 42469. .43278
/note="L1P92 repeat: matches 5315. .6154 of consensus"
repeat_region 44183. .44258
/note="L1M4 repeat: matches 2886. .2965 of consensus"
repeat_region 44917. .45185
/note="AluSg repeat: matches 1. .295 of consensus"
repeat_region 46431. .46800
/note="L1PA13 repeat: matches 5787. .6156 of consensus"
repeat_region 46819. .46945
/note="HERV1 repeat: matches 397. .1283 of consensus"
repeat_region 46959. .47269
/note="AluX repeat: matches 1. .311 of consensus"
repeat_region 49450. .50350
/note="HERV1 repeat: matches 397. .1283 of consensus"
repeat_region 50724. .50941
/note="L1M4 repeat: matches 2184. .2413 of consensus"
repeat_region 50956. .51095
/note="4 copies 35 mer 73% conserved"
repeat_region 50985. .51032
/note="L2 copies 4 mer tata 79% conserved"
repeat_region 51015. .51122
/note="4 copies 27 mer 72% conserved"
repeat_region 51358. .51503
/note="L1M4 repeat: matches 2422. .2571 of consensus"
repeat_region 51491. .51668
/note="L1M61 repeat: matches 4563. .4733 of consensus"
repeat_region 51675. .51852
/note="AluSg repeat: matches 1. .179 of consensus"
repeat_region 51853. .52151
/note="L1M61 repeat: matches 4894. .5201 of consensus"
repeat_region 52194. .52911
/note="L1M61 repeat: matches 5284. .6017 of consensus"
repeat_region 54539. .54764
/note="L2 repeat: matches 2514. .2750 of consensus"

Query Match 99.8%; Score 1036.4; DB 9; Length 141028;
Best Local Similarity 99.9%; Pred. No. 1e-222; Indels 0; Gaps 0;
Matches 1037; Conservative 1; Indels 0; Gaps 0;

1 TCAGGAATGATGCCCTTTGGCCACATATATTAATTTCTGTGTGAAAAACAATGG 60
132175 TCAGGAATGATGCCCTTTGGCCACATATTAATTTCTGTGTGAAAAACAATGG 132116

61 TCAAAATGATGCCGCTTCCCTGTACAGTTTAATAGTGCCTAATAATTGTACACACATC 120
132115 TCAAAATGATGCCGCTTCCCTGTACAGTTTAATAGTGCCTAATAATTGTACACACATC 132056

121 GTTGCATATGATATGTTATGTTCTATATACACTTCAACCACTTCATACCCCAACA 180
132056 GTTGCATATGATATGTTATGTTCTATATACACTTCAACCACTTCATACCCCAACA 131996

181 AATTGGCTCATTCATTCATGCGCACTGTGACCTTCTTGGGGTGTGTCATGCT 240
131995 AATTGGCTCATTCATTCATGCGCACTGTGACCTTCTTGGGGTGTGTCATGCT 131936

241 TACAGTATGCTAGATTCGCTGACGACCTGTGATTTTGGAGAAATCTTCTGTAAAT 300
131935 TACAGTATGCTAGATTCGCTGACGACCTGTGATTTTGGAGAAATCTTCTGTAAAT 131876

301 CACACAAGACCGAATATGCTGAGCTCAGGCTCATTTTTCATTTGCTTCAATCTCC 360
131875 CACACAAGACCGAATATGCTGAGCTCAGGCTCATTTTTCATTTGCTTCAATCTCC 131816

361 ATTGAACCGCTACTATGCTGTGTGTATCATCATGAGATATTAAGCAAGATGAAATCTTG 420

Db 131815 ATTACCGCTACTATGCTGTGTGATCCATGAGATATTAAGCCAAAGTAATATCTTG 131756
Qy 421 GTTATTTTGTGATGATCTTATTAAGTGAAGTGTCCCTGCTGTTTTGCAATTTGGAATG 480
Db 131755 GTTATTTTGTGATGATCTTATTAAGTGAAGTGTCCCTGCTGTTTTGCAATTTGGAATG 131696
Qy 481 ATCTTTCGAGCTAACTTCAAGGCGCTGAAGAGATATTAACAACATGTTTCACTGC 540
Db 131695 ATCTTTCGAGCTAACTTCAAGGCGCTGAAGAGATATTAACAACATGTTTCACTGC 131636
Qy 541 AGAGAGAGTGTCTGTCTCTCTTTAGCAAAATATCTGGGCTACTGACCTTTATGACTTCT 600
Db 131635 AGAGAGAGTGTCTGTCTCTCTTTAGCAAAATATCTGGGCTACTGACCTTTATGACTTCT 131576
Qy 601 TTTTATATACCTGATCTATTAATGATATGCTATGCTATTAACAGATATATCTTATGCTTAA 660
Db 131575 TTTTATATACCTGATCTATTAATGATATGCTATGCTATTAACAGATATATCTTATGCTTAA 131516
Qy 661 GAAACAGCAAGATTAAATTAAGTATGCCAATCAGAACTCCAAATTGGATTGAAATGAA 720
Db 131515 GAAACAGCAAGATTAAATTAAGTATGCCAATCAGAACTCCAAATTGGATTGAAATGAA 131456
Qy 721 AATGGAATTTCAAGACCAAGAAAGAAAGCTGTGAAGACATTGGGGAATTGTATGGGA 780
Db 131455 AATGGAATTTCAAGACCAAGAAAGAAAGCTGTGAAGACATTGGGGAATTGTATGGGA 131396
Qy 781 GTTTTCTATATATGCTGTGTGCCCTTCTTATCTGTACAGTATGAGACCTTTCTTCA 840
Db 131395 GTTTTCTATATATGCTGTGTGCCCTTCTTATCTGTACAGTATGAGACCTTTCTTCA 131336
Qy 841 TACATATATTCACCTTACTTGAATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 131335 TACATATATTCACCTTACTTGAATGATGATGATGATGATGATGATGATGATGATGATG 131276
Qy 901 TTTTATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 131275 TTTTATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 131216
Qy 961 CTGTTTGTGTAATTTTCCAAAAGATTGATCAGGTGTAAATTTATTTTGGAAATGAGT 1020
Db 131215 CTGTTTGTGTAATTTTCCAAAAGATTGATCAGGTGTAAATTTATTTTGGAAATGAGT 131156
Qy 1021 TCATGAATTAATTAATTT 1038
Db 131155 TCATGAATTAATTAATTT 131138

RESULT 3
AX921196 1039 bp DNA linear PAT 18-DEC-2003
LOCUS AX921196
DEFINITION Sequence 189 from Patent WO02068652.
ACCESSION AX921196
VERSION AX921196.1 GI:40214830
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Nov-x proteins and nucleic acids encoding same
JOURNAL Patent: WO 02068652-A 189 06-SEP-2002;
FEATURES
LOCATION/Qualifiers
1. 1039
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.5%; Score 1032.4; DB 6; Length 1039;
Best Local Similarity 99.9%; Pred. No. 1.5e-231;
Matches 1033; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GAATGATGCCCTTTTGCCCAATATTAATTAATTTCCGTGTGAAGAAAGCACTGGTCA 64
Db 1 GAATGATGCCCTTTTGCCCAATATTAATTAATTTCCGTGTGAAGAAAGCACTGGTCA 60
Qy 65 ATGATGTCCTGCTTCCCTGTACAGTTTAATGAGTCAATTAATTCGACCACTGCTTG 124
Db 61 ATGATGTCCTGCTTCCCTGTACAGTTTAATGAGTCAATTAATTCGACCACTGCTTG 120
Qy 125 GCATCTGATAGTATGATGTTTCTATATACACTTCAACCACTCACTCCCAACAAATT 184
Db 121 GCATCTGATAGTATGATGTTTCTATATACACTTCAACCACTCACTCCCAACAAATT 180
Qy 185 GGCCTATTCATTCATGAGCCACTGTGAATTTCTTGTGGGGGTGTCTGGTCACTGCTTCA 244
Db 181 GGCCTATTCATTCATGAGCCACTGTGAATTTCTTGTGGGGGTGTCTGGTCACTGCTTCA 240
Qy 245 GTATGATGATATCTGCTGAGACACTGTTGTATTTTGGAGAAAGTCTTGTAAATTCACA 304
Db 241 GTATGATGATATCTGCTGAGACACTGTTGTATTTTGGAGAAAGTCTTGTAAATTCACA 300
Qy 305 CAAGCACCAGATTAATGCTGAGCTCAGCTTCATTTTCCATTTGCTTTCATCTCCATG 364
Db 301 CAAGCACCAGATTAATGCTGAGCTCAGCTTCATTTTCCATTTGCTTTCATCTCCATG 360
Qy 365 ACCGCTACTATGCTGTGTGTGATCCACTGAGATTAAGCCAAAGTAAATATCTTGTTA 424
Db 361 ACCGCTACTATGCTGTGTGTGATCCACTGAGATTAAGCCAAAGTAAATATCTTGTTA 420
Qy 425 TTTGTGATGATATCTTATAGTATGATGATGATGATGATGATGATGATGATGATGATG 484
Db 421 TTTGTGATGATATCTTATAGTATGATGATGATGATGATGATGATGATGATGATGATG 480
Qy 485 TTTGTGAGCTTAACCTCAAGGCGCTGAAGAGATATTAACAACATGTTCACTGAGAG 544
Db 481 TTTGTGAGCTTAACCTCAAGGCGCTGAAGAGATATTAACAACATGTTCACTGAGAG 540
Qy 545 GAGGTCTCTGCTTCTTTTGAAGAAATATCTGGGATACAGCTTTATGACTTCTTTT 604
Db 541 GAGGTCTCTGCTTCTTTTGAAGAAATATCTGGGATACAGCTTTATGACTTCTTTT 600
Qy 605 ATATACCTGATCTATTAATGATATGATGATGATGATGATGATGATGATGATGATG 664
Db 601 ATATACCTGATCTATTAATGATATGATGATGATGATGATGATGATGATGATGATG 660
Qy 665 AGCAAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 724
Db 661 AGCAAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 725 GAATTCACAAAGCAAGAAAGAAAGCTGTGAAGACATTGGGGATGATGAGAGTTT 784
Db 721 GAATTCACAAAGCAAGAAAGAAAGCTGTGAAGACATTGGGGATGATGAGAGTTT 780
Qy 785 TCTTAATATGCTGTGCTCTTCTTATATGCTGACGTCAAGACCTTTTCTTCACTACA 844
Db 781 TCTTAATATGCTGTGCTCTTCTTATATGCTGACGTCAAGACCTTTTCTTCACTACA 840
Qy 845 TTAATTCACCTTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 904
Db 841 TTAATTCACCTTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Qy 905 ATCCAAATGTTTATGATTTTCTATCCCTGTTTGAAGAAAGCACTGAAGATGATGCTGT 964
Db 901 ATCCAAATGTTTATGATTTTCTATCCCTGTTTGAAGAAAGCACTGAAGATGATGCTGT 960
Qy 965 TTGTGTAATTTTCCAAAAGATTATATCAGGTGTAAATTTATTTTGGAAATGAGTCA 1024
Db 961 TTGTGTAATTTTCCAAAAGATTATATCAGGTGTAAATTTATTTTGGAAATGAGTCA 1020
Qy 1025 AGAATTAATTAATTT 1038
Db 1021 AGAATTAATTAATTT 1034

```
RESULT 4
AX921198      1039 bp   DNA      linear   PAT 18-DEC-2003
LOCUS         Sequence 191 from Patent WO02068652.
ACCESSION    AX921198
VERSION      AX921198.1 GI:40214831
KEYWORDS
SOURCE
ORGANISM     Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS      Nov-x proteins and nucleic acids encoding same
TITLE        Patent: WO 02068652-A 191 06-SEP-2002;
JOURNAL      Location/Qualifiers
FEATURES
             source
             1..1039
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"

ORIGIN
Query Match      99.5%; Score 1032.4; DB 6; Length 1039;
Best Local Similarity 99.9%; Pred. No. 1.5e-231;
Matches 1033; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 GAATGATGCCCTTTTGGCCCAATATATATATTTCTGTGTGAAAAACAACCTGCTCA 64
DB      1 GAATGATGCCCTTTTGGCCCAATATATATATTTCTGTGTGAAAAACAACCTGCTCA 60
QY      65 ATGATGTCGGTCTTCCCTGTACAGTTATAGTGTGCTATATTTGACACACCTGCTG 124
DB      61 ATGATGTCGGTCTTCCCTGTACAGTTATAGTGTGCTATATTTGACACACCTGCTG 120
QY      125 GCAATCTGATGTTATTTGTTCTATATACACCTTCAACCACTTCATACCCCACAATT 184
DB      121 GCAATCTGATGTTATTTGTTCTATATACACCTTCAACCACTTCATACCCCACAATT 180
QY      185 GGGCTATTCATTCATGCGCACTGCGCACTTTCTTCTGGGGGTCTGCGTCACTGCTTACA 244
DB      181 GGGCTATTCATTCATGCGCACTGCGCACTTTCTTCTGGGGGTCTGCGTCACTGCTTACA 240
QY      245 GTATGATGATCTGCTGAGCACTGTGGTATTTTGGAGAAGCTTCTGTAAATTCACA 304
DB      241 GTATGATGATCTGCTGAGCACTGTGGTATTTTGGAGAAGCTTCTGTAAATTCACA 300
QY      305 CAAGCACCGACATTAATGCTGAGCTCAGCTCCATTTTCCATTTGCTTCATCTCATTTG 364
DB      301 CAAGCACCGACATTAATGCTGAGCTCAGCTCCATTTTCCATTTGCTTCATCTCATTTG 360
QY      365 ACCGCTACTATGCTGTGTGTATCCACTGATATTAAGCCAAAGATGATATCTTGTTA 424
DB      361 ACCGCTACTATGCTGTGTGTATCCACTGATATTAAGCCAAAGATGATATCTTGTTA 420
QY      425 TTTGTGTGATGATCTTCAATTAAGTGGAGTGTCCCTGCTGTTTGGATTTGGATGATCT 484
DB      421 TTTGTGTGATGATCTTCAATTAAGTGGAGTGTCCCTGCTGTTTGGATTTGGATGATCT 480
QY      485 TTCTGAGCTAAACTTCAAGGCGCTGAAGAGATATATTCAACAACATGTTCACTGAGAG 544
DB      481 TTCTGAGCTAAACTTCAAGGCGCTGAAGAGATATATTCAACAACATGTTCACTGAGAG 540
QY      545 GAGGTGTGCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTAAGACTTCTTTT 604
DB      541 GAGGTGTGCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTAAGACTTCTTTT 600
QY      605 ATATACCTGATCTATATGTTATGTTGCTATATACGAATATATCTTATCGCTAAAGAC 664
DB      601 ATATACCTGATCTATATGTTATGTTGCTATATACGAATATATCTTATCGCTAAAGAC 660
QY      665 AGGCAAGATTAATTAAGTATGCAATCAGAAGCTCCAAATGGAGTTGAAATGAAAAATG 724
DB      661 AGGCAAGATTAATTAAGTATGCAATCAGAAGCTCCAAATGGAGTTGAAATGAAAAATG 720
```

```
DB      661 AGGCAAGATTAATTAAGTATGCAATCAGAAGCTCCAAATGGAGTTGAAATGAAAAATG 720
QY      725 GAATTTGCAAAAGCAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGATGGAGTTT 784
DB      721 GAATTTGCAAAAGCAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGATGGAGTTT 780
QY      785 TCTTAATATGCTGTGTCCTTCTTATCTGTACAGTCATGGAACCTTTCTTCACTACA 844
DB      781 TCTTAATATGCTGTGTCCTTCTTATCTGTACAGTCATGGAACCTTTCTTCACTACA 840
QY      845 TTATTTCCACCTTATGGAATGATGTGTGATTTGGTGTGCTACTTGAACCTTACATTTA 904
DB      841 TTATTTCCACCTTATGGAATGATGTGTGATTTGGTGTGCTACTTGAACCTTACATTTA 900
QY      905 ATCCAAATGTTATGATATTTTCTATCTGCTGTTAGAAAAGCACTGAAGATGATGCTGT 964
DB      901 ATCCAAATGTTATGATATTTTCTATCTGCTGTTAGAAAAGCACTGAAGATGATGCTGT 960
QY      965 TTGTTAAATTTTCCAAAAGATTCATCCAGGTGTAATTTATTTTGGAAATGAGTTCA 1024
DB      961 TTGTTAAATTTTCCAAAAGATTCATCCAGGTGTAATTTATTTTGGAAATGAGTTCA 1020
QY      1025 AGAATTATTAATTT 1038
DB      1021 AGAATTATTAATTT 1034

RESULT 5
AX921206      1039 bp   DNA      linear   PAT 18-DEC-2003
LOCUS         Sequence 199 from Patent WO02068652.
ACCESSION    AX921206
VERSION      AX921206.1 GI:40214835
KEYWORDS
SOURCE
ORGANISM     Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS      Nov-x proteins and nucleic acids encoding same
TITLE        Patent: WO 02068652-A 199 06-SEP-2002;
JOURNAL      Location/Qualifiers
FEATURES
             source
             1..1039
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"

ORIGIN
Query Match      99.5%; Score 1032.4; DB 6; Length 1039;
Best Local Similarity 99.9%; Pred. No. 1.5e-231;
Matches 1033; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 GAATGATGCCCTTTTGGCCCAATATATATATTTCTGTGTGAAAAACAACCTGCTCA 64
DB      1 GAATGATGCCCTTTTGGCCCAATATATATATTTCTGTGTGAAAAACAACCTGCTCA 60
QY      65 ATGATGTCGGTCTTCCCTGTACAGTTATAGTGTGCTATATTTGACACACCTGCTG 124
DB      61 ATGATGTCGGTCTTCCCTGTACAGTTATAGTGTGCTATATTTGACACACCTGCTG 120
QY      125 GCAATCTGATGTTATTTGTTCTATATACACCTTCAACCACTTCATACCCCACAATT 184
DB      121 GCAATCTGATGTTATTTGTTCTATATACACCTTCAACCACTTCATACCCCACAATT 180
QY      185 GGGCTATTCATTCATGCGCACTGCGCACTTTCTTCTGGGGGTCTGCGTCACTGCTTACA 244
DB      181 GGGCTATTCATTCATGCGCACTGCGCACTTTCTTCTGGGGGTCTGCGTCACTGCTTACA 240
QY      245 GTATGATGATCTGCTGAGCACTGTGGTATTTTGGAGAAGCTTCTGTAAATTCACA 304
DB      241 GTATGATGATCTGCTGAGCACTGTGGTATTTTGGAGAAGCTTCTGTAAATTCACA 300
```

OY		305	CAAGACCGGACAATTATGCTGAGCTCAGCCCTCATTTTGCATTTGTCTTGATCTTCATTG	367
Db		301	CAMGACACCGACATTATNGCTGAGCTCAGCCCTCATTTTGCATTTGTCTTGATCTTCATTG	367
OY		365	ACCGETACTANGCTGTGTGATCCATCGAGATATAAAGCCAAAGATGAATCTTGTTTA	422
Db		361	ACCGETACTANGCTGTGTGATCCATCGAGATATAAAGCCAAAGATGAATCTTGTTTA	422
OY		425	TTTGTGTATGATCTTCACTTAGTTGAGTGTCCCTGCTGTTTTTGGATTGGATGATCT	484
Db		421	TTTGTGTATGATCTTCACTTAGTTGAGTGTCCCTGCTGTTTTTGGATTGGATGATCT	484
OY		485	TTTCGAGCTTAACTTCCAAAAGCCCTGAAGAATATATTAACAACATGTTCACTGCAAG	544
Db		481	TTTCGAGCTTAACTTCCAAAAGCCCTGAAGAATATATTAACAACATGTTCACTGCAAG	544
OY		545	GAGGTGCTCTGTCTTCTTTAGCAAAAATATCTGGGTACTGACCTTATGACTCTTTT	604
Db		541	GAGGTGCTCTGTCTTCTTTAGCAAAAATATCTGGGTACTGACCTTATGACTCTTTT	600
OY		605	ATATACCTGGATCTCATATATGTTATGNTGCTATTAACAGATATATCTTATGCTPAAAGAC	664
Db		601	ATATACCTGGATCTCATATATGTTATGNTGCTATTAACAGATATATCTTATGCTPAAAGAC	660
OY		665	AGGCAAGATTAAATTAGTATGCCAATCAGAAGCTCCAAATTTGGATTGGAAAAATG	724
Db		661	AGGCAAGATTAAATTAGTATGCCAATCAGAAGCTCCAAATTTGGATTGGAAAAATG	720
OY		725	GAATTTCCAAACCAAGAAAGAAAGAAAGCTGTGAAGACATTTGGAGATTGTATGGAGTTT	784
Db		721	GAATTTCCAAACCAAGAAAGAAAGAAAGCTGTGAAGACATTTGGAGATTGTATGGAGTTT	780
OY		785	TCCTAATATGCTGTGATCCCTTTCTTATCTGTACAGTCATGACCCCTTTTCTCATACA	844
Db		781	TCCTAATATGCTGTGATCCCTTTCTTATCTGTACAGTCATGACCCCTTTTCTCATACA	840
OY		845	TTATTTCCACCTATTTTGAATGATGTGTGATTTGGTTGGCTACCTTGAACCTCATATTA	904
Db		841	TTATTTCCACCTATTTTGAATGATGTGTGATTTGGTTGGCTACCTTGAACCTCATATTA	900
OY		905	ATCCAATGTTTATGACATTTTTCATCCCTGGTTTAAAGAACATGAAAGATGATGCTGT	964
Db		901	ATCCAATGTTTATGACATTTTTCATCCCTGGTTTAAAGAACATGAAAGATGATGCTGT	960
OY		965	TTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAAATATTTTGGATTGAGTTCAT	1022
Db		961	TTGGTAAATTTTCCAAAAGATTCATCCAGGTGTAAATATTTTGGATTGAGTTCAT	1020
OY		1025	AGATTATTAATATT 1038	
Db		1021	AGATTATTAATATT 1034	
RESULT 6				
LOCUS	AX497139	1125 bp	DNA	linear PAT 26-SEP-2002
DEFINITION	Sequence 1 from Patent WO022801.			
ACCESSION	AX497139			
VERSION	AX497139.1 GI:23342534			
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
JOURNAL	Bunzow,J.R., Grandy,D.K. and Sondere,M.			
FEATURES	Mammalian receptor genes and uses			
source	Patent: WO 022801-A 1 21-MAR-2002;			
	Oregon Health & Science University (US)			
	location/Qualifiers			
	1..1125			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			

```

CDS
    /db_xref="taxon:9606"
    21..1040
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAD48641.1"
    /db_xref="GI:23342535"
    /translation="MMPFCHNI INISCVANNNSNDVRLSLVWLITLTLLVGNLIV
    IYVISHKQLHTPTMWLISHMAVDPLGLCVMSVMSRLEHCYRPREVCKHTS
    DMLSSAFSLFHSFISIDRYVAVCDPLKAKMNIYVLCWNI RISMSPVPAFGMIR
    LEMFKGAEIYKHVHCRGGGLVFPFSKISGULTMSPFYLPDSIMLCVYRIYALN
    EQRRLSDANQQLQGLBEKRNKGISQSKRKAVKTLGIWGFLLCWCFPLCTVWDP
    LHSIDPITLNDVLWFGYLNSTFENMVAFYFVPMWRKALKMFLPKI FQKDSRCKLF
    LELSS"

ORIGIN
Query Match      99.1%; Score 1028.4; DB 6; Length 1125;
Best Local Similarity 99.4%; Pred No.1.3e-230;
Matches 1032; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 TCAGGATATATGCGCTTTTGCCACAATAATTAATTAATTTCTGTGTGAAAAACAACCTGG 60
Db      15 TCAGGATATATGCGCTTTTGCCACAATAATTAATTAATTTCTGTGTGAAAAACAACCTGG 74

QY      61 TCAAATGATGTCCGTCCTCCCTGACAGTTAATNGTCTCAATATTTGCAACACTC 120
Db      75 TCAAATGATGTCCGTCCTCCCTGACAGTTAATNGTCTCAATATTTGCAACACTC 134

QY      121 GTTGGCAATCTGATAGTATTGTTTCTAATATCAACTTCAACAACACTTCATACCCACAA 180
Db      135 GTTGGCAATCTGATAGTATTGTTTCTAATATCAACTTCAACAACACTTCATACCCACAA 194

QY      181 AATTGACCATTCATTCATGAGCCACAGTGAGACTTCTTCTGGGGGTCTGTGATGCTT 240
Db      195 AATTGACCATTCATTCATGAGCCACAGTGAGACTTCTTCTGGGGGTCTGTGATGCTT 254

QY      241 TACAGTATGTGAAGATCTGTGAGACACTGTGTGATTTTGGAGAGCTTCTGTGAAAAAT 300
Db      255 TACAGTATGTGAAGATCTGTGAGACACTGTGTGATTTTGGAGAGCTTCTGTGAAAAAT 314

QY      301 CACACAAACACCGACATTAATGCTGAGCTGACGCCCTCCATTTTCCATTGTCTTCATCTCC 360
Db      315 CACACAAACACCGACATTAATGCTGAGCTGACGCCCTCCATTTTCCATTGTCTTCATCTCC 374

QY      361 ATTGACCGCTAATAAGCTGTGTGTGATCCACTGAGATATATAACCAAGATGAATATCTTG 420
Db      375 ATTGACCGCTAATAAGCTGTGTGTGATCCACTGAGATATATAACCAAGATGAATATCTTG 434

QY      421 GTTATTGTGTGATGATCTTCATTTAGTTGGAAGTGTCCCTGTCTGTTTGGATTTGGAAATG 480
Db      435 GTTATTGTGTGATGATCTTCATTTAGTTGGAAGTGTCCCTGTCTGTTTGGATTTGGAAATG 494

QY      481 ATCTTTCTGAGCTTAACCTCAAGGCGCTGAGAGATATATTAACAAATGTGTACATCGC 540
Db      495 ATCTTTCTGAGCTTAACCTCAAGGCGCTGAGAGATATATTAACAAATGTGTACATCGC 554

QY      541 AGAGGAGGTGCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCT 600
Db      555 AGAGGAGGTGCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCT 614

QY      601 TTTTATATACCTGATCTATATATGATATGTCATTAAGATATATCTTAATTCGTCTAA 660
Db      615 TTTTATATACCTGATCTATATATGATATGTCATTAAGATATATCTTAATTCGTCTAA 674

QY      661 GAACAGGCAAGATTATTATGTGTGCCAATCAAGAGCTTCAAAATTGATTTGGAAATGAAA 720
Db      675 GAACAGGCAAGATTATTATGTGTGCCAATCAAGAGCTTCAAAATTGATTTGGAAATGAAA 734

QY      721 AATGAAATTTCACAAAGCAAAGAGAAAGCTGTGAGAGACATTTGGGATTTGTATGGGA 780
Db      735 AATGAAATTTCACAAAGCAAAGAGAAAGCTGTGAGAGACATTTGGGATTTGTATGGGA 794

QY      781 GTTTTCCTAATATGCTGTGCGCTTTCTTATCTGTACAGTCATGAGACCTTTTCTTAC 840

```

Db 795 GTTTCCTATATAGCTGTGCCCCCTTCTTATCTGTACAGTCAGACCCCTTCTTAC 854
Qy 841 TACATTAATCCACCTACTTGAATGATGTTGTTGGCTTGAACCTTACA 900
Db 855 TCAATTAATCCACCTACTTGAATGATGTTGTTGGCTTGAACCTTACA 914
Qy 901 TTTAATCCAAATGTTATGATCTTCTTCTGTTTGAAGACATGAAATGATG 960
Db 915 TTTAATCCAAATGTTATGATCTTCTTCTGTTTGAAGACATGAAATGATG 974
Qy 961 CTGTTGGTAAATTTTCCAAAAAGATTCATCCAGGTGTAATTTATTTTGAATGAT 1020
Db 975 CTGTTGGTAAATTTTCCAAAAAGATTCATCCAGGTGTAATTTATTTTGAATGAT 1034
Qy 1021 TCATGAATTAATTAAT 1038
Db 1035 TCATGAATTAATTAAT 1052

RESULT 7
AX147844 1023 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 89 from Patent WO0136473.
ACCESSION AX147844
VERSION AX147844.1 GI:14346843
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
Vogeli, G., Wood, L.S., Parodi, L.A., Hiesbach, R.R., Lind, P.,
Slichtom, J., Schellin, K.A., Kayles, P.S., Bannigan, C.M., Ruff, V.,
Sejitz, T., and Huff, R.M.
Novel 9 protein-coupled receptors
Patent: WO 0136473-A 89 25-MAY-2001;
JOURNAL PHARMACIA & UPJOHN COMPANY (US)
FEATURES
source 1..1023
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.4%; Score 1021.4; DB 6; Length 1023;
Best Local Similarity 99.9%; Pred. No. 5.7e-229;
Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GGAATGATGCCCTTTGGCCAAATATAATTAATTTCCGTGTGAAAAACAATGCTCA 63
Db 1 GGAATGATGCCCTTTGGCCAAATATAATTAATTTCCGTGTGAAAAACAATGCTCA 60
Qy 64 AATGATGTCGTCCTTCCCTGTACAGTTTAATGTCCTAATAATTCGACACATCGTT 123
Db 61 AATGATGTCGTCCTTCCCTGTACAGTTTAATGTCCTAATAATTCGACACATCGTT 120
Qy 124 GGCATCTGATAGTTAATTTCTATATACACTTGAACAATTATATCCCAAAAT 183
Db 121 GGCATCTGATAGTTAATTTCTATATACACTTGAACAATTATATCCCAAAAT 180
Qy 184 TGGCTCATTCATTCATGCGCACTGTGACTTCTTCTGGGGTGTCTGGTCAATGCTTAC 243
Db 181 TGGCTCATTCATTCATGCGCACTGTGACTTCTTCTGGGGTGTCTGGTCAATGCTTAC 240
Qy 244 AGTATGATGATCTGCTGAGCACTGTGGTATTTTGGAGAGTCTTCTGTAATTTAC 303
Db 241 AGTATGATGATCTGCTGAGCACTGTGGTATTTTGGAGAGTCTTCTGTAATTTAC 300
Qy 304 ACAAGACCGACATTAATGCTGAGCTCAGCTCCATTTTCATTTGCTTCAATCTCAT 363
Db 301 ACAAGACCGACATTAATGCTGAGCTCAGCTCCATTTTCATTTGCTTCAATCTCAT 360
Qy 364 GACCGCTACTATGCTGTGTGATCCACTGAGATATAAGCCAAAGATGAATATCTTGTT 423

Db 361 GACCGCTACTATGCTGTGTGATCCACTGAGATATAAGCCAAAGATGAATATCTTGTT 420
Qy 424 ATTGTGTGATGATCTTCAATTAATGATGAGTGTCCCTGCTGTTTGGATTTGGAATGATC 483
Db 421 ATTGTGTGATGATCTTCAATTAATGATGAGTGTCCCTGCTGTTTGGATTTGGAATGATC 480
Qy 484 TTTCTGAGCTTAACTTCAAAAGCCCTGAGAGATATATTACAACATGTTCACTGCAGA 543
Db 481 TTTCTGAGCTTAACTTCAAAAGCCCTGAGAGATATATTACAACATGTTCACTGCAGA 540
Qy 544 GGAGGTGCTGTCTTCTTTAGCAAAATATCTGGGGTATCTGACCTTTATGACTTCTTTT 603
Db 541 GGAGGTGCTGTCTTCTTTAGCAAAATATCTGGGGTATCTGACCTTTATGACTTCTTTT 600
Qy 604 TATATACCTGATCTATTAATGTTATGATGCTATTAAGAAATATATCTTATGCTTAAAGA 663
Db 601 TATATACCTGATCTATTAATGTTATGATGCTATTAAGAAATATATCTTATGCTTAAAGA 660
Qy 664 CAGCAAGATTAATAGTATGATCCATCAAGAGCTCAAAATTTGATTTGGAATGAATAAT 723
Db 661 CAGCAAGATTAATAGTATGATCCATCAAGAGCTCAAAATTTGATTTGGAATGAATAAT 720
Qy 724 GGAATTTTCAAAAGCAAAAGAAAGAAAGCTGGAAGCAATTTGGGATTTGTATGGAGTT 783
Db 721 GGAATTTTCAAAAGCAAAAGAAAGAAAGCTGGAAGCAATTTGGGATTTGTATGGAGTT 780
Qy 784 TTCTTAATATGCTGTGCCCCCTTCTTATCTGTACAGTCATGAGACCTTTTCTACTTC 843
Db 781 TTCTTAATATGCTGTGCCCCCTTCTTATCTGTACAGTCATGAGACCTTTTCTACTTC 840
Qy 844 ATTATTTCCACCTACTTGAATGATGTTGTTGTTGGCTACTGTAATCTACATTT 903
Db 841 ATTATTTCCACCTACTTGAATGATGTTGTTGTTGGCTACTGTAATCTACATTT 900
Qy 904 AATCCAAATGTTATGATCTTTTCTATCTGTTTGAAGAACATGAAATGATCTG 963
Db 901 AATCCAAATGTTATGATCTTTTCTATCTGTTTGAAGAACATGAAATGATCTG 960
Qy 964 TTTGGTAAATTTTCCAAAAAGATTCATCCAGGTGTAATTTTGGAAATTTGAGTTCA 1023
Db 961 TTTGGTAAATTTTCCAAAAAGATTCATCCAGGTGTAATTTTGGAAATTTGAGTTCA 1020
Qy 1024 TAG 1026
Db 1021 TAG 1023

RESULT 8
AX521893 1023 bp DNA linear PAT 24-OCT-2002
LOCUS AX521893
DEFINITION Sequence 89 from Patent WO02064789.
ACCESSION AX521893
VERSION AX521893.1 GI:24410795
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
Lind, P., Parodi, L.A., Vogeli, G., and Wood, L.S.
G protein-coupled receptor
Patent: WO 02064789-A 89 22-AUG-2002;
JOURNAL PHARMACIA & UPJOHN COMPANY (US)
FEATURES
source 1..1023
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.4%; Score 1021.4; DB 6; Length 1023;
Best Local Similarity 99.9%; Pred. No. 5.7e-229;

Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 4 GGAATGATGCGCTTTGGCACAATATTAATATTTCTGTGTGAAAAACAATGCTCA 63
Db 1 GGAATGATGCGCTTTGGCACAATATTAATATTTCTGTGTGAAAAACAATGCTCA 60
QY 64 AATGATGCGCTGCTTCCCTGACAGTTAAATGAGTCTATATTTGACACAGCTGTT 123
Db 61 AATGATGCGCTGCTTCCCTGACAGTTAAATGAGTCTATATTTGACACAGCTGTT 120
QY 124 GGCATCTGATGATGTTATGTTCTATATCACTTCAACAACCTTCAATCCCAACAAT 183
Db 121 GGCATCTGATGATGTTATGTTCTATATCACTTCAACAACCTTCAATCCCAACAAT 180
QY 184 TGGCTCATTCATTCATCCAGCAGTGTGACCTTTCTTGGGGGTGCTGGTCAAGCTTAC 243
Db 181 TGGCTCATTCATTCATCCAGCAGTGTGACCTTTCTTGGGGGTGCTGGTCAAGCTTAC 240
QY 244 AGTATGATGATGATCTGCTGAGCACTGTGTGATTTTGAAGAAGCTTCTGTAAATTCAC 303
Db 241 AGTATGATGATGATCTGCTGAGCACTGTGTGATTTTGAAGAAGCTTCTGTAAATTCAC 300
QY 304 ACAAGCACCGACATTTATGCTGAGCTCAGCCCAATTTTCAATTTGCTTCAATCTCAAT 363
Db 301 ACAAGCACCGACATTTATGCTGAGCTCAGCCCAATTTTCAATTTGCTTCAATCTCAAT 360
QY 364 GACCGCTACTATGCTGTGTGTATCACTGAGATATTAAGCAAGATGAATATCTTGTT 423
Db 361 GACCGCTACTATGCTGTGTGTATCACTGAGATATTAAGCAAGATGAATATCTTGTT 420
QY 424 ATTTGTGATGATCTTCAATTTAGTGGAGTCCCTGCTGCTTTTGCATTTGGAATGATC 483
Db 421 ATTTGTGATGATCTTCAATTTAGTGGAGTCCCTGCTGCTTTTGCATTTGGAATGATC 480
QY 484 TTTTCTGAGCTAACTTCAAAAGCGCTGAAGAGATATTTAACAACATGTTCACTGACGA 543
Db 481 TTTTCTGAGCTAACTTCAAAAGCGCTGAAGAGATATTTAACAACATGTTCACTGACGA 540
QY 544 GGAAGTGTGCTGTCTTCTTTAGCAAAATATCTGGGGTCTGACCTTTATGACCTTTCTTT 603
Db 541 GGAAGTGTGCTGTCTTCTTTAGCAAAATATCTGGGGTCTGACCTTTATGACCTTTCTTT 600
QY 604 TATATCTGATCTATATATGTTATGTTCTATTAACAGAAATATATCTTATGCTTAAAGA 663
Db 601 TATATCTGATCTATATATGTTATGTTCTATTAACAGAAATATATCTTATGCTTAAAGA 660
QY 664 CAGGCAAGTAAATTAAGTATGCCAATCAAGAGCTTCAAAATGGAATGGAATGAAAAAT 723
Db 661 CAGGCAAGTAAATTAAGTATGCCAATCAAGAGCTTCAAAATGGAATGGAATGAAAAAT 720
QY 724 GGAATTTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGGATGTTGATGGAGTT 783
Db 721 GGAATTTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGGATGTTGATGGAGTT 780
QY 784 TTCCTAATATGCTGTGCTTCTTTATCTGTACAGTATGACCTTTCTTCACTAC 843
Db 781 TTCCTAATATGCTGTGCTTCTTTATCTGTACAGTATGACCTTTCTTCACTAC 840
QY 844 ATTATTTCACTACTTGAATGATGTGTGATTTGTTGGTCTTCTGAACTCTACATTT 903
Db 841 ATTATTTCACTACTTGAATGATGTGTGATTTGTTGGTCTTCTGAACTCTACATTT 900
QY 904 AATCAATGTTTATGCAATTTTCTATCTTGTGTTAAGAAAGCACTGAAGATGATCTG 963
Db 901 AATCAATGTTTATGCAATTTTCTATCTTGTGTTAAGAAAGCACTGAAGATGATCTG 960
QY 964 TTTGTGTAATTTTCAAAAAAGATTCACAGGTGTAATTTATTTTGAATGAGTTCA 1023
Db 961 TTTGTGTAATTTTCAAAAAAGATTCACAGGTGTAATTTATTTTGAATGAGTTCA 1020
QY 1024 TAG 1026
Db 1021 TAG 1023
```

```
RESULT 9
LOCUS CO879788 1020 bp DNA linear PAT 11-OCT-2004
DEFINITION Sequence 1 from Patent WO2004083851.
ACCESSION CO879788
VERSION CO879788.1 GI:54033726
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Goliz, S., Brueggemeier, U. and Geerts, A.
Diagnosics and therapeutics for diseases associated with trace
amine receptor 1 (ta1)
Patent: WO 2004083851-A 1 30-SEP-2004;
Bayer Healthcare AG (DE)
FEATURES
source
1..1020
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
```

Query Match 98.3%; Score 1020; DB 6; Length 1020;
Best Local Similarity 100.0%; Pred. No. 1,2e-228;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 7 ATGATGCGCTTTGGCACAATATTAATATTTCCGTGTGAAAAACAATGCTCAAT 66
Db 1 ATGATGCGCTTTGGCACAATATTAATATTAATATTTCCGTGTGAAAAACAATGCTCAAT 60
QY 67 GATGTCGGTCTTCCCTGTAACAGTTTAATGAGTCTATATTTCTGACCACTGCTGGGC 126
Db 61 GATGTCGGTCTTCCCTGTAACAGTTTAATGAGTCTATATTTCTGACCACTGCTGGGC 120
QY 127 ATCTGATGATTTTCTATATCACTTCAACAACCTTCAATCCCAACAATTTGG 186
Db 121 ATCTGATGATTTTCTATATCACTTCAACAACCTTCAATCCCAACAATTTGG 180
QY 187 CTCATTCATTCATGCGCACTGTGACCTTCTTGGGGGTGCTGGTATGCTTACAGT 246
Db 181 CTCATTCATTCATGCGCACTGTGACCTTCTTGGGGGTGCTGGTATGCTTACAGT 240
QY 247 ATGTGAGATCTGCTGAGCACTGTTGTATTTTGAAGAAGCTTCTGTAAATTCACACA 306
Db 241 ATGTGAGATCTGCTGAGCACTGTTGTATTTTGAAGAAGCTTCTGTAAATTCACACA 300
QY 307 AGCACCGACATTAATGCTGAGCTCAGCTCCATTTTCAATTTGCTTTCATCTCCATGAC 366
Db 301 AGCACCGACATTAATGCTGAGCTCAGCTCCATTTTCAATTTGCTTTCATCTCCATGAC 360
QY 367 CGCTACTATGCTGTGTGATCACTGAGATATTAAGCAAGATGAATATCTTGTTAT 426
Db 361 CGCTACTATGCTGTGTGATCACTGAGATATTAAGCAAGATGAATATCTTGTTAT 420
QY 427 TGTGTGATGATCTTCAATAGTTGAGTGTCCCTGCTGTTTTCATTTGCAATTTGATCTTT 486
Db 421 TGTGTGATGATCTTCAATAGTTGAGTGTCCCTGCTGTTTTCATTTGCAATTTGATCTTT 480
QY 487 CTGAGCTAAACTTCAAAAGCGCTGAAGAGATATTAACAACATGTTCACTGACAGGA 546
Db 481 CTGAGCTAAACTTCAAAAGCGCTGAAGAGATATTAACAACATGTTCACTGACAGGA 540
QY 547 GGTGCTGTGCTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCTTTTAT 606
Db 541 GGTGCTGTGCTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCTTTTAT 600
QY 607 ATACTGATCTATATGTTATGTTCTATTTAACAATATATCTTATGCTTAAAGAACAG 666
Db 601 ATACTGATCTATATGTTATGTTCTATTTAACAATATATCTTATGCTTAAAGAACAG 660
```


QY 667 GCAGATTAAATAGTATGATCCAAATGAGAGTCCAAATTTGATGGAATGAAAAATGGA 726
Db 661 GCAGATTAAATAGTATGATCCAAATGAGAGTCCAAATTTGATGGAATGAAAAATGGA 720
QY 727 ATTTACAAAGCAAGAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTGATGGAGTTTTC 786
Db 721 ATTTACAAAGCAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTGATGGAGTTTTC 780
QY 787 CTAAATAGTGTGTCCTTTCTTTATCTGTACAGTCATGAGACCTTTTCTTACACTACAT 846
Db 781 CTAAATAGTGTGTCCTTTCTTTATCTGTACAGTCATGAGACCTTTTCTTACACTACAT 840
QY 847 ATTTCACTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906
Db 841 ATTTCACTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 907 CCAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
Db 901 CCAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 967 GGTAAATTTTCCAAAGAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTGATGG 1026
Db 961 GGTAAATTTTCCAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTGATGGATGAT 1020

RESULT 10
AX549353
LOCUS AX549353 1020 bp DNA linear PAT 26-NOV-2002
DEFINITION Sequence 638 from Patent WO2061087.
ACCESSION AX549353
VERSION AX549353.1 GI:25813987
KEYWORDS

SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS

TITLE

1 Bumer, G.C., Roub, C.L. and Brown, J.P.
Antigenic peptides, such as for G protein-coupled receptors
(GPCRs), antibodies thereto, and systems for identifying such
antigenic peptides
Patent: WO 02061087-A 638 08-AUG-2002;

JOURNAL

Lifespan Biosciences, Inc. (US)

FEATURES

source

1.1020
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 98.3%; Score 1020; DB 6; Length 1020;
Best Local Similarity 100.0%; Pred. No. 1.2e-228;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 ATGATGCCCTTTGGCAACAAT 66
Db 1 ATGATGCCCTTTGGCAACAAT 60
QY 67 GATGTCCTGCTTCCCTGTACAGTTAATGATGATGATGATGATGATGATGATGATGATGAT 126
Db 61 GATGTCCTGCTTCCCTGTACAGTTAATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 127 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 186
Db 121 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 187 CTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 246
Db 181 CTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 240
QY 247 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 306

Db 241 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 307 AGCACCGCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 366
Db 301 AGCACCGCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 367 CGCTACTATGCTGT 426
Db 361 CGCTACTATGCTGT 420
QY 427 TGT 486
Db 421 TGT 480
QY 487 CTGAGCTTAACTTCAAGAGGCTGTGAAGATATATATATATATATATATATATATATATAT 546
Db 481 CTGAGCTTAACTTCAAGAGGCTGTGAAGATATATATATATATATATATATATATATATAT 540
QY 547 GGTGCTCTGCTCTTCTTTAGCAAAATATCTGGGATCTGACCTTTATGATCTTTTAT 606
Db 541 GGTGCTCTGCTCTTCTTTAGCAAAATATCTGGGATCTGACCTTTATGATCTTTTAT 600
QY 607 ATACTGATCTAT 666
Db 601 ATACTGATCTAT 660
QY 667 GCAGATTAAATAGTATGATCCAAATGAGAGTCCAAATTTGATGGAATGAAAAATGGA 726
Db 661 GCAGATTAAATAGTATGATCCAAATGAGAGTCCAAATTTGATGGAATGAAAAATGGA 720
QY 727 ATTTACAAAGCAAGAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTGATGGAGTTTTC 786
Db 721 ATTTACAAAGCAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTGATGGAGTTTTC 780
QY 787 CTAAATAGTGTGTCCTTTCTTTATCTGTACAGTCATGAGACCTTTTCTTACACTACAT 846
Db 781 CTAAATAGTGTGTCCTTTCTTTATCTGTACAGTCATGAGACCTTTTCTTACACTACAT 840
QY 847 ATTTCACTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 906
Db 841 ATTTCACTACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 907 CCAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
Db 901 CCAATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 967 GGTAAATTTTCCAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTGATGGATGAT 1026
Db 961 GGTAAATTTTCCAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTGATGGATGAT 1020

RESULT 11

AF380185

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AF380185 1020 bp DNA linear PRI 01-AUG-2001
AF380185 Homo sapiens trace amine receptor 1 (TA1) gene, complete cde.
AF380185.1 GI:14600073
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1020)
Boroweky, B., Adam, N., Jones, K.A., Raddecz, R., Artyushyn, R.,
Ogozalek, K.L., Durkin, M.M., Lakhiani, P.P., Bonini, J.A.,
Pachirana, S., Boyle, N., Pu, X., Kouranov, E., Lichblau, H.,
Ochoa, F.Y., Branchek, T.A. and Gerald, C.
Trace amines: identification of a family of mammalian G
protein-coupled receptors
Proc. Natl. Acad. Sci. U.S.A. 98 (16), 8966-8971 (2001)
21374364
11459929
2 (bases 1 to 1020)

AUTHORS Adham, N., Raddatz, R., Artyushyn, R., Durkin, M.M., Ogozalek, K.L., Lakhiani, P.P., Kouzanova, E.V., Ochoa, Y., Boyle, N., Pu, X., Branchek, T., Gerald, C., Blackburn, T.P., and Borowsky, B.

TITLE Pharmacological characterization of a G protein-coupled receptor, Tal, activated by trace amines and amphetamines: Evidence for major species differences in drug profile and tissue distribution

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1020)

AUTHORS Borowsky, B., Ogozalek, K.L., Jones, K.A. and Adham, N.

TITLE Direct Submission

JOURNAL Submitted (12-MAY-2001) Target Discovery & Assessment, Synaptic Pharmaceutical Corp., 215 College Road, Paramus, NY 07652, USA

FEATURES

Source

1..1020

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="6"

/map="6q23.2"

<1..>1020

/gene="TA1"

<1..>1020

/gene="TA1"

/product="TA1"

1..1020

/product="trace amine receptor 1"

/gene="TA1"

/note="G protein-coupled receptor"

/codon_start=1

/product="trace amine receptor 1"

/protein_id="AAK71236.1"

/db_xref="GI:1460074"

/translation="MMPCHNINISCVNNNSNDVRASLYSLMVLILITLVGNLIVYSISHPKQIAHTPTWMLHSMATVDPLGLVMPYSMSASHCWPGVPECKHISTDMLSSASIFHLSPISIDRYAVCPRLRYKAMNLIICVMFISMSPVAPAFMITLELNFGAEIYYKHVHCRGCGSVSPSKISGLTETMTEPVIQSGIMLCVYRRIYIAKBOARLISDNOKLQIGLEKNGISOSKSKRAKVTGIWGFPLICWCPFFICTWDPFLHRTIPPLINDVLIMFGYILNSTFNPVTAFFYPMFRKALKMMLFKRIFQKSSRCKLFLKLS"

ORIGIN

Query Match 98.3%; Score 1020; DB 9; Length 1020;

Best Local Similarity 100.0%; Pred. No. 1.2e-228; Indels 0; Gaps 0;

Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGATGCCCTTTGGCACAATATAATTAATTTCTGTGTAAGAAACAAGTGTCAAT 66

DB 1 ATGATGCCCTTTGGCACAATATAATTAATTTCTGTGTAAGAAACAAGTGTCAAT 60

QY 67 GATGCCGCTCCCTGACAGTTAATGCTCATTAATTTCTGACCACTGCTGGC 126

DB 61 GATGCCGCTCCCTGACAGTTAATGCTCATTAATTTCTGACCACTGCTGGC 120

QY 127 AATCGATAGTATGTTTCTATATCACTCAACAACCTCATACCCCAACAATGG 186

DB 121 AATCGATAGTATGTTTCTATATCACTCAACAACCTCATACCCCAACAATGG 180

QY 187 CTGATTCATTCATGCGCACTGTGCACTTTCTTGGGGTGTCTGTCATGCTTACAGT 246

DB 181 CTGATTCATTCATGCGCACTGTGCACTTTCTTGGGGTGTCTGTCATGCTTACAGT 240

QY 247 AAGGTGAGATCTGCGGACGACGTTGGTATTTGGAGAAGCTTCTGTAATTTCAACA 306

DB 241 ATGGGAGATCTGCGGACGACGTTGGTATTTGGAGAAGCTTCTGTAATTTCAACA 300

QY 307 AGCACCGACATTATGCTAGCTCAGCCTCATTTTCCATTTGTCTTCATCTCCATTGAC 366

DB 301 AGCACCGACATTATGCTAGCTCAGCCTCATTTTCCATTTGTCTTCATCTCCATTGAC 360

QY 367 CGCTACTATGCTGTGTGTGATCCACTGAGATATAAGCCAAAGATGAATATCTGTTATT 426

DB 361 CGCTACTATGCTGTGTGTGATCCACTGAGATATAAGCCAAAGATGAATATCTGTTATT 420

QY 427 TGTGTGATGATCTTCATTAGTTGAGAGTGTCCCTGCTGTTTGTGCAATTTGAATATCTTT 486

DB 421 TGTGTGATGATCTTCATTAGTTGAGAGTGTCCCTGCTGTTTGTGCAATTTGAATATCTTT 480

QY 487 CTGAGCTAAACTTCMAAGCGCTGAGAGATATATTCAACAATGTTCACTGCAGAGGA 546

DB 481 CTGAGCTAAACTTCMAAGCGCTGAGAGATATATTCAACAATGTTCACTGCAGAGGA 540

QY 547 GGTGCTCTGTCTTTCTTTAGCAAAATATCTGGGGTACTGACCTTATGACTTCTTTTAT 606

DB 541 GGTGCTCTGTCTTTCTTTAGCAAAATATCTGGGGTACTGACCTTATGACTTCTTTTAT 600

QY 607 ATACCTGATCTATATGTTATGTTGTCTATTACAGAAATATCTTATGCTTAAAGAACAG 666

DB 601 ATACCTGATCTATATGTTATGTTGTCTATTACAGAAATATCTTATGCTTAAAGAACAG 660

QY 667 GCAGATTAATTAATGATGCCAATCAGAGCTCCAAATTTGATTTGAATGAAAAATGGA 726

DB 661 GCAGATTAATTAATGATGCCAATCAGAGCTCCAAATTTGATTTGAATGAAAAATGGA 720

QY 727 ATTTCAAGAAAGAAAGAAAGAGCTGTGAAGACATTGGGGATTGTGATGGAGTTTC 786

DB 721 ATTTCAAGAAAGAAAGAAAGAGCTGTGAAGACATTGGGGATTGTGATGGAGTTTC 780

QY 787 CTAATATGCTGTGCCCCCTTTCTTTATCTGTACAGTCAATGACCCCTTTCTTCACTACAT 846

DB 781 CTAATATGCTGTGCCCCCTTTCTTTATCTGTACAGTCAATGACCCCTTTCTTCACTACAT 840

QY 847 ATTCACCTACTTGTGAATGATGTGTGATTTGTTGGCTAATTGAACCTCACTTAAT 906

DB 841 ATTCACCTACTTGTGAATGATGTGTGATTTGTTGGCTAATTGAACCTCACTTAAT 900

QY 907 CCAATGGTTATAGATTTTCTATCTGTTGTTGAAGAAACAGCTGAAGATGATGCTGTT 966

DB 901 CCAATGGTTATAGATTTTCTATCTGTTGTTGAAGAAACAGCTGAAGATGATGCTGTT 960

QY 967 GGTAAAAATTTCCAAAAAGATTCAATCCAGGTGTAATTAATTTTGAATGATGATGATG 1026

DB 961 GGTAAAAATTTCCAAAAAGATTCAATCCAGGTGTAATTAATTTTGAATGATGATGATG 1020

RESULT 12

AX498188

LOCUS AX498188 1020 bp DNA linear PAT 26-SEP-2002

DEFINITION Sequence 11 from Patent WO0242461.

ACCESSION AX498188

VERSION AX498188.1 GI:23343115

KEYWORDS

SOURCE

ORGANISM

synthetic construct

synthetic construct

other sequences; artificial sequences.

REFERENCE

1

AUTHORS Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.

TITLE Endogenous and non-endogenous versions of human g protein-coupled receptors

JOURNAL Patent: WO 0242461-A 11 30-MAY-2002;

ARENA Pharmaceuticals, Inc. (US)

FEATURES

Source

1..1020

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Novel Sequence"

ORIGIN

Query Match 98.1%; Score 1018.4; DB 6; Length 1020;

Best Local Similarity 99.9%; Pred. No. 2.9e-228; Indels 0; Gaps 0;

Matches 1019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ATGATGCCCTTTGGCACAATATAATTAATTTCTGTGTAAGAAACAAGTGTCAAT 66

DB 1 ATGATGCCCTTTGGCACAATATAATTAATTTCTGTGTAAGAAACAAGTGTCAAT 60

QY 67 GATGCCGCTCCCTGACAGTTAATGCTCATTAATTTCTGACCACTGCTGGC 126

Db 61 GATGTCGGGCTTCCCTGTACAGTTAAAGTGCTCATTAATCTGCACACACTCGTTGGC 120

Qy 127 AATCGATGATGATATGTTCTATATACAGCTTCAAAACAATTCAATCCCAACAATGG 186

Db 121 AATCGATGATGATATGTTCTATATACAGCTTCAAAACAATTCAATCCCAACAATGG 180

Qy 187 CTCATTCATTCATGAGCCACCTGTGACCTTCTTGGGGGTGTCTGTACATGCTTACAGT 246

Db 181 CTCATTCATTCATGAGCCACCTGTGACCTTCTTGGGGGTGTCTGTACATGCTTACAGT 240

Qy 247 ATGGTGAGATCTGTGAGACAGCTGTTGGATTTTGGAGAAAGCTTCTGTAATAATTCACACA 306

Db 241 ATGGTGAGATCTGTGAGACAGCTGTTGGATTTTGGAGAAAGCTTCTGTAATAATTCACACA 300

Qy 307 AGCACCAGCATTAATGCTGAGCTCAGCCTCCATTTCCATTTCTTTCATCTCCCAATGAC 366

Db 301 AGCACCAGCATTAATGCTGAGCTCAGCCTCCATTTCCATTTCTTTCATCTCCCAATGAC 360

Qy 367 CGCTACTATGCTGTGTGTATGATCCACTGAGATATTAAGCCAGATGAAATCTTGGTTAT 426

Db 361 CGCTACTATGCTGTGTGTATGATCCACTGAGATATTAAGCCAGATGAAATCTTGGTTAT 420

Qy 427 TGTGTGATGATCTTCATTAAGTTGAGAGTGTCCCTGCTGTTTGGCATTTGGAATGATCTTT 486

Db 421 TGTGTGATGATCTTCATTAAGTTGAGAGTGTCCCTGCTGTTTGGCATTTGGAATGATCTTT 480

Qy 487 CTGGAGCTTAACTTCAAAAGCCGTGAGAGATATTAACAACATCTTACCTGACAGAGA 546

Db 481 CTGGAGCTTAACTTCAAAAGCCGTGAGAGATATTAACAACATCTTACCTGACAGAGA 540

Qy 547 GGTGCTCTGTCTTCTTGAAGAAATATCTGGGGTACTGACCTTATGACTCTTTTAT 606

Db 541 GGTGCTCTGTCTTCTTGAAGAAATATCTGGGGTACTGACCTTATGACTCTTTTAT 600

Qy 607 ATACTCGATCTATTATGTTATGTGTCTATTAACAGATATATCTTATGCTTAAAGAACG 666

Db 601 ATACTCGATCTATTATGTTATGTGTCTATTAACAGATATATCTTATGCTTAAAGAACG 660

Qy 667 GCAAAATTAATTAAGTGAAGCCCAATGAGCTCCAAATTTGGAATGGAATGAATGA 726

Db 661 GCAAAATTAATTAAGTGAAGCCCAATGAGCTCCAAATTTGGAATGGAATGAATGA 720

Qy 727 ATTTCAAAAGCAAAAGAAAGAGCTGTGAAGACATTTGGGATTTGTATGGGAGTTTC 786

Db 721 ATTTCAAAAGCAAAAGAAAGAGCTGTGAAGACATTTGGGATTTGTATGGGAGTTTC 780

Qy 787 CTAAATATGCTGTGCCCTTCTTATCTGTACAGTCATGAGACCTTTTCTTCACTACAT 846

Db 781 CTAAATATGCTGTGCCCTTCTTATCTGTACAGTCATGAGACCTTTTCTTCACTACAT 840

Qy 847 ATTTCAACCTACTTTGAATGATGTGTGATTTGGTCTACTTGAACCTTCACTATTAAT 906

Db 841 ATTTCAACCTACTTTGAATGATGTGTGATTTGGTCTACTTGAACCTTCACTATTAAT 900

Qy 907 CCAATGATTTATGACATTTTCTATCTTGTGTTAGAAAGACATGAAGATGATGCTGTT 966

Db 901 CCAATGATTTATGACATTTTCTATCTTGTGTTAGAAAGACATGAAGATGATGCTGTT 960

Qy 967 GGTAAATTTTCCAAAAGATTCATCAGGTGTAAATTAATTTTGGAAATGAGTTGATNG 1026

Db 961 GGTAAATTTTCCAAAAGATTCATCAGGTGTAAATTAATTTTGGAAATGAGTTGATNG 1020

RESULT 13

AF200627 1020 bp DNA linear PRI 10-JAN-2002

LOCUS AF200627 Homo sapiens putative catecholamine receptor gene, complete cds.

DEFINITION AF200627

ACCESSION AF200627.1 GI:10441576

VERSION

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS

Bunzow, J.R., Sonders, M.S., Artavanisoul, S., Harrison, L.M., Zhang, G., Quigley, D.I., Darland, T., Suchland, K.L., Pasumamula, S., Kennedy, J.L., Olson, S.B., Magenle, R.B., Amara, S.G., and Grandy, D.K. Amphetamine, 3,4-methylenedioxymethamphetamine, lysergic acid diethylamide, and metabolites of the catecholamine neurotransmitters are agonists of a rat trace amine receptor Mol. Pharmacol. 60 (6), 1181-1188 (2001)

JOURNAL

MEDLINE

21580235

11723224

2 (bases 1 to 1020)

REFERENCE

Bunzow, J.R. and Grandy, D.K.

AUTHORS

TITLE

Direct Submision

JOURNAL

Submitted (22-SEP-1999) Physiology and Pharmacology, Oregon Health Sciences University, 3181 S.W. Sam Jackson Park Rd., Portland, OR 97201, USA

FEATURES

source

1..1020

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>1020

/product="putative catecholamine receptor"

1..1020

/codon_start=1

/product="putative catecholamine receptor"

/protein_id="AA017112.1"

/db_xref="gi:10441577"

/translation="MPECHNIINISCVKNNMSNDVRASLYSLVLIILTLVGLNLIIVISISHPKQLHTPTMLIHSMAVDPLFGCLVMPVSRASLHCVFSEVFEKISTDIMSASIFHLSPISIDRYAVDCLPRKKNMLIVCNMIFISMSPAVAFAMIFLEINPKGAEIYVGHVHRCGGSVSPFSISVLTPMPSFYI BGSIMLCVYRIYLIAR EEARLISPAKQLOIGLEMRKGISRSKRKAVKTLGIYMGFLICMCPFICTWDPF LHYIIPPLNDVLIWFGYLNSTFENMVAFYIPWERRKALKNMLFKKIRFOKSSRCKLFLELS"

ORIGIN

Query Match 98.1%; Score 1018.4; DB 9; Length 1020;

Best Local Similarity 99.9%; Pred. No. 2,9e-226;

Matches 1019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ATGATGCCCTTTGGCAAAATATTAATATTTCTGTGTGAAGAAACAATGTCAAAT 66

Db 1 ATGATGCCCTTTGGCAAAATATTAATATTTCTGTGTGAAGAAACAATGTCAAAT 60

Qy 67 GATGTCGGTGTCTCCCTGTACAGTTTAAAGTGCTCATTAATTTCTGACCACTGCTGGC 126

Db 61 GATGTCGGTGTCTCCCTGTACAGTTTAAAGTGCTCATTAATTTCTGACCACTGCTGGC 120

Qy 127 AATCGATGATGATATGTTCTTATATACAGCTTCAAAACAATTCAATCCCAACAATGG 186

Db 121 AATCGATGATGATATGTTCTTATATACAGCTTCAAAACAATTCAATCCCAACAATGG 180

Qy 187 CTCATTCATTCATGAGCCACCTGTGACCTTCTTGGGGGTGTCTGTACATGCTTACAGT 246

Db 181 CTCATTCATTCATGAGCCACCTGTGACCTTCTTGGGGGTGTCTGTACATGCTTACAGT 240

Qy 247 ATGGTGAGATCTGTGAGACAGCTGTTGGATTTTGGAGAAAGCTTCTGTAATAATTCACACA 306

Db 241 ATGGTGAGATCTGTGAGACAGCTGTTGGATTTTGGAGAAAGCTTCTGTAATAATTCACACA 300

Qy 307 AGCACCAGCATTAATGCTGAGCTCAGCCTCCATTTCCATTTCTTTCATCTCCCAATGAC 366

Db 301 AGCACCAGCATTAATGCTGAGCTCAGCCTCCATTTCCATTTCTTTCATCTCCCAATGAC 360

Qy 367 CGCTACTATGCTGTGTGTATGATCCACTGAGATATTAAGCCAGATGAAATCTTGGTTAT 426

Db 361 CGCTACTATGCTGTGTGTATGATCCACTGAGATATTAAGCCAGATGAAATCTTGGTTAT 420

Qy 427 TGTGTGATGATCTTCATTAAGTTGAGAGTGTCCCTGCTGTTTGGCATTTGGAATGATCTTT 486

Db 421 TGTGTGATGATCTTCAATAGTGAAGTGCCTGCTGTTTTCATTTGCAATGATGATCTTT 480
Qy 487 CTGAGAGTAACTTTCAAAAGCGCTGAAGAGATATATTACAAATGTTCACTGACAGGA 546
Db 481 CTGAGAGTAACTTTCAAAAGCGCTGAAGAGATATATTACAAATGTTCACTGACAGGA 540
Qy 547 GGTGCTCTGCTCTTTAGCAAAATATCTGGGTAAGTCACTTTTGTGACTTCTTTTAT 606
Db 541 GGTGCTCTGCTCTTTAGCAAAATATCTGGGTAAGTCACTTTTGTGACTTCTTTTAT 600
Qy 607 ATTCCTGATCTATTTATGTTATGTCATATACAGATATATCTTATCGCTAAAGACAG 666
Db 601 ATACCTGATCTATTTATGTTATGTCATATACAGATATATCTTATCGCTAAAGACAG 660
Qy 667 GCAAGATTATATGATGATGCAATGCAAGAGCTCAAAATGGAATGGAATGGAATGGA 726
Db 661 GCAAGATTATATGATGATGCAATGCAAGAGCTCAAAATGGAATGGAATGGAATGGA 720
Qy 727 ATTTCACAAAGCAAAAGAAAGAAAGCTGTGAAGACATGCGGATGTTGATGGAGTTTC 786
Db 721 ATTTCACAAAGCAAAAGAAAGAAAGCTGTGAAGACATGCGGATGTTGATGGAGTTTC 780
Qy 787 CTAAATATGCTGTCCTCTTTCTTTATCTGTACAGTATGACCTTTTCTTCACTACAT 846
Db 781 CTAAATATGCTGTCCTCTTTCTTTATCTGTACAGTATGACCTTTTCTTCACTACAT 840
Qy 847 ATTTCACCTACTTGTGAATGATGTTGATTTGTTGGTCTTGAACCTTCACTTATAT 906
Db 841 ATTTCACCTACTTGTGAATGATGTTGATTTGTTGGTCTTGAACCTTCACTTATAT 900
Qy 907 CCAATGTTTATGATCTTTTCTATCTGTTAGAAAGCACTGAAGATGATGCTGTTT 966
Db 901 CCAATGTTTATGATCTTTTCTATCTGTTAGAAAGCACTGAAGATGATGCTGTTT 960
Qy 967 GGTAAATTTTCCAAAAGATCATCCAGCTGTAATTTTGTGAAATGAGTTCAATG 1026
Db 961 GGTAAATTTTCCAAAAGATCATCCAGCTGTAATTTTGTGAAATGAGTTCAATG 1020

RESULT 14
AY180374 1020 bp mRNA linear PRI 19-JAN-2003
LOCUS AY180374
DEFINITION Homo sapiens trace amine receptor 1 mRNA, complete cds.
ACCESSION AY180374
VERSION AY180374.1 GI:27803881
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Kopatz,S.A., Aronstam,R.S. and Sharma,S.V.
TITLE Isolation of cDNA coding for Human Trace Amine Receptor (TAR) 1
JOURNAL Unpublished
AUTHORS Kopatz,S.A., Aronstam,R.S. and Sharma,S.V.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2002) Guthrie cDNA Resource Center, Guthrie
Research Institute, 1 Guthrie Square, Sayre, PA 18840, USA
FEATURES
SOURCE
1..1020
/organism="Homo sapiens"
/mol_type="mRNA"
1..1020
/note="G-protein coupled receptor"
/codon_start=1
/product="Trace amine receptor 1"
/protein_id="AA022154.1"
/db_xref="GI:27803882"
/translation="IMPFCHNIINISCVKNMSNDVRAASHLSLVLIILITIVGNLIV
IVSISHPKOLHTPTNMLIHMAITVDPLGLGLVMPYSVRAEHHWYRGEGVCKHTST
DIMSSASIFHLISIDRYAVCDPLRIYAKKMLIIVICMIPISMSVPAVFAFMIF

ORIGIN
Query Match 98.1%; Score 1018.4; DB 9; Length 1020;
Best Local Similarity 99.9%; Pred. No. 2.9e-228;
Matches 1019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
LELNFKGABEYIKHVHRCGCSVFESKISGLVLFMTISFYIPGSLMLCVYRIYLIAK
EQARLISDANKLQJGEMKNGISQSKERKAVKLIQIVMGVFLCMCPFLICTVNDPF
LHYIIPPTLNDVLIMFGYLNSTFNDMVYAFYPMWRKALKNMLFKIKIPOKSSRCKLF
LELS"

Qy 7 ATGATGCCCTTTGGCACAATATATATATTTCTGTGTGAAAAACAAGCTGCAAT 66
Db 1 ATGATGCCCTTTGGCACAATATATATATTTCTGTGTGAAAAACAAGCTGCAAT 60
Qy 67 GATGTCGCTGCTCCCTGTACAGTTTAAATGCTCATATTTGACCACTGCTGGC 126
Db 61 GATGTCGCTGCTCCCTGTACAGTTTAAATGCTCATATTTGACCACTGCTGGC 120
Qy 127 AATCATAGTTATTTGTTCTTATATACACTTCAACACTTCAACCTTCAACCTTCA 186
Db 121 AATCATAGTTATTTGTTCTTATATACACTTCAACACTTCAACCTTCAACCTTCA 180
Qy 187 CTCATTCAATCCAGGCACTGTGGAATTTCTGTGGGGTGTCTGATCATGCTTACAGT 246
Db 181 CTCATTCAATCCAGGCACTGTGGAATTTCTGTGGGGTGTCTGATCATGCTTACAGT 240
Qy 247 ATGATGAGATCTGCTGACACTGTGTTGTTTGGAGAAGCTTCTGTAAATTCACACA 306
Db 241 ATGATGAGATCTGCTGACACTGTGTTGTTTGGAGAAGCTTCTGTAAATTCACACA 300
Qy 307 AGCACCGACATTTAGCTGAGCTGAGCTGAGCTTTCATTTGCTTTCATTCATTCATTC 366
Db 301 AGCACCGACATTTAGCTGAGCTGAGCTGAGCTTTCATTTGCTTTCATTCATTCATTC 360
Qy 367 CGTACTATGCTGTGTGTATCCACTGAGATATTAAGCCAAAGATGAATATCTGTTATT 426
Db 361 CGTACTATGCTGTGTGTATCCACTGAGATATTAAGCCAAAGATGAATATCTGTTATT 420
Qy 427 TGTGTGATGATCTTCAATAGTGAAGTGCCTGCTGTTTGTGATTTGGAATGATCTTT 486
Db 421 TGTGTGATGATCTTCAATAGTGAAGTGCCTGCTGTTTGTGATTTGGAATGATCTTT 480
Qy 487 CTGAGCTAAACCTTCAAGGCGCTGAAGAGATATATTACAAACATGTTCACTGACAGGA 546
Db 481 CTGAGCTAAACCTTCAAGGCGCTGAAGAGATATATTACAAACATGTTCACTGACAGGA 540
Qy 547 GGTGCTCTGCTCTTTTACCAAAATATCTGGGTAAGTCACTTTTATGACTTCTTTTAT 606
Db 541 GGTGCTCTGCTCTTTTACCAAAATATCTGGGTAAGTCACTTTTATGACTTCTTTTAT 600
Qy 607 ATACCTGATCTATTTATGTTATGTCATATACAGATATATCTTATCGCTAAAGACAG 666
Db 601 ATACCTGATCTATTTATGTTATGTCATATACAGATATATCTTATCGCTAAAGACAG 660
Qy 667 GCAAGATTATATGATGATGCAATGCAAGAGCTCAAAATGGAATGGAATGGAATGGA 726
Db 661 GCAAGATTATATGATGATGCAATGCAAGAGCTCAAAATGGAATGGAATGGAATGGA 720
Qy 727 ATTTCACAAAGCAAAAGAAAGAAAGCTGTGAAGACATGCGGATGTTGATGGAGTTTC 786
Db 721 ATTTCACAAAGCAAAAGAAAGAAAGCTGTGAAGACATGCGGATGTTGATGGAGTTTC 780
Qy 787 CTAAATATGCTGTCCTCTTTCTTTATCTGTACAGTATGACCTTTTCTTCACTACAT 846
Db 781 CTAAATATGCTGTCCTCTTTCTTTATCTGTACAGTATGACCTTTTCTTCACTACAT 840
Qy 847 ATTTCACCTACTTGTGAATGATGTTGATTTGTTGGTCTTGAACCTTCACTTATAT 906
Db 841 ATTTCACCTACTTGTGAATGATGTTGATTTGTTGGTCTTGAACCTTCACTTATAT 900
Qy 907 CCAATGTTTATGATCTTTTCTATCTGTTTGAAGAAAGCACTGAAGATGATGCTGTTT 966
Db 901 CCAATGTTTATGATCTTTTCTATCTGTTTGAAGAAAGCACTGAAGATGATGCTGTTT 960

QY 967 GGTAAATTTTCCAAAAGATTCAGGTGTAATTTTGGAAATGAGTTCATAG 1026
DB 961 GGTAAATTTTCCAAAAGATTCAGGTGTAATTTTGGAAATGAGTTCATAG 1020

RESULT 15
LOCUS COJ39824 1017 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 25758 from Patent WO02068579.
ACCESSION COJ39824
VERSION COJ39824.1 GI:42343845
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS 1
TITLE Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kite, such as nucleic acid arrays, comprising a majority of
humans or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 25758 06-SEP-2002;
PE Corporation

FEATURES
source
1..1017
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 97.8%; Score 1015.4; DB 6; Length 1017;
Best Local Similarity 99.9%; Pred. No. 1.5e-227;
Matches 1016; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 ATGCCCTTTTGGCCAAATATTAATTAATTTCTGTGAAAAACAAGTGTCAATGAT 69
DB 1 ATGCCCTTTTGGCCAAATATTAATTAATTTCTGTGAAAAACAAGTGTCAATGAT 60

QY 70 GTCCGGTCTTCCCTGTACAGTTAAAGTGTCTAATTTGACCACTCGTGGCAAT 129
DB 61 GTCCGGTCTTCCCTGTACAGTTAAAGTGTCTAATTTGACCACTCGTGGCAAT 120

QY 130 CTGATAGTATTTCTATATCACTCAACCACTCATACCCCAACAATTGGCTC 189
DB 121 CTGATAGTATTTCTATATCACTCAACCACTCATACCCCAACAATTGGCTC 180

QY 190 ATTCAATTCATGACCACTGTGACCTTCTGTGGGGTGTCTGTCAATGCTTACAGTATG 249
DB 181 ATTCAATTCATGACCACTGTGACCTTCTGTGGGGTGTCTGTCAATGCTTACAGTATG 240

QY 250 GTGAGATCTGCTGAGACACTGTGTAATTTTGAAGAAGTCTTCTGTAAATTCACACAAAGC 309
DB 241 GTGAGATCTGCTGAGACACTGTGTAATTTTGAAGAAGTCTTCTGTAAATTCACACAAAGC 300

QY 310 ACCGACATTAATGCTGAGCTCAGGCTCATTTTTCATTTGCTTTCATCTCAATGACCGC 369
DB 301 ACCGACATTAATGCTGAGCTCAGGCTCATTTTTCATTTTGTCTTTCATCTCAATGACCGC 360

QY 370 TACTATGCTGTGTGATCACTGAGATATAAGCCAAAGATGAATCTGTGTAATTTGT 429
DB 361 TACTATGCTGTGTGATCACTGAGATATAAGCCAAAGATGAATCTGTGTAATTTGT 420

QY 430 GTGATGATCTTCAATTAATGAGTGTCCCTGTCTTTTTCATTTGGAATGATCTTTCTG 489
DB 421 GTGATGATCTTCAATTAATGAGTGTCCCTGTCTTTTTCATTTGGAATGATCTTTCTG 480

QY 490 GAGCTAACTTCAAAAGCCCTGAGAGATATAATTAACAATGTTCACTGACAGAGAGGT 549
DB 481 GAGCTAACTTCAAAAGCCCTGAGAGATATAATTAACAATGTTCACTGACAGAGAGGT 540

QY 550 TGCTCTGCTTTTGAACAATAATATCTGAGGATGACCTTTATGACCTTTTATATA 609
DB 541 TGCTCTGCTTTTGAACAATAATATCTGAGGATGACCTTTATGACCTTTTATATA 600

DB 541 TGCTCTGCTTTTGAACAATAATATCTGAGGATGACCTTTATGACCTTTTATATA 600
QY 610 CCTGATCTAATATGTAATGTAATGTAATTAACAATATATCTTATGCTTAAAGACAGCA 669
DB 601 CCTGATCTAATATGTAATGTAATGTAATTAACAATATATCTTATGCTTAAAGACAGCA 660

QY 670 AGATTAAATTAATGATGCAATCAGAAAGCTCCAAATTTGAAATGAAATGAAATGAAAT 729
DB 661 AGATTAAATTAATGATGCAATCAGAAAGCTCCAAATTTGAAATGAAATGAAATGAAAT 720

QY 730 TCACAAAGCAAAAGAAAGAGCTGTGAAGACATTTGGGATTTGATGGAAGTTTTCCTA 789
DB 721 TCACAAAGCAAAAGAAAGAGCTGTGAAGACATTTGGGATTTGATGGAAGTTTTCCTA 780

QY 790 ATATGCTGTGCTCTTTCTTATCTGTACAGTACATGACCTTTCTTCACTACATTAAT 849
DB 781 ATATGCTGTGCTCTTTCTTATCTGTACAGTACATGACCTTTCTTCACTACATTAAT 840

QY 850 CCACCTACTTTGAATGATGTGTGATTTGGTTGGCTACTGAACTTCACTACATTAATCA 909
DB 841 CCACCTACTTTGAATGATGTGTGATTTGGTTGGCTACTGAACTTCACTACATTAATCA 900

QY 910 ATGCTTTATGCAATTTTCTTATCTGTGTTTGAAGACACTGAAGATGATGCTTTTGGT 969
DB 901 ATGCTTTATGCAATTTTCTTATCTGTGTTTGAAGACACTGAAGATGATGCTTTTGGT 960

QY 970 AAAATTTTCCAAAAGATTCAGGTGTAATTAATTTTGAATGAGTTCATAG 1026
DB 961 AAAATTTTCCAAAAGATTCAGGTGTAATTAATTTTGAATGAGTTCATAG 1017

Search completed: February 12, 2005, 01:47:21
Job time : 3284 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 00:39:32 ; Search time 463 Seconds
(without alignments)
13271.472 Million cell updates/sec

Title: US-09-980-145-5
Perfect score: 1038
Sequence: 1 tcaggaatgacgaccttctg.....gtccatagatattatatt 1038

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_16Dec04: *
1: geneeqn19808: *
2: geneeqn19908: *
3: geneeqn20008: *
4: geneeqn20018: *
5: geneeqn20018: *
6: geneeqn20028: *
7: geneeqn20028: *
8: geneeqn20038: *
9: geneeqn20038: *
10: geneeqn20038: *
11: geneeqn20038: *
12: geneeqn20048: *
13: geneeqn20048: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1038	100.0	1038	4 AAC90247	AAC90247 Human SNO
2	1036.4	99.8	3501	2 AAH26920	AAH26920 Human G P
3	1032.4	99.5	1039	6 ADH48915	ADH48915 NOV84B co
4	1032.4	99.5	1039	6 ADH48905	ADH48905 NOV81A co
5	1032.4	99.5	1039	6 ADH48907	ADH48907 NOV81B co
6	1028.4	99.1	1125	6 AB158316	AB158316 Human tra
7	1021.4	98.4	1023	4 AAH51013	AAH51013 Human ncp
8	1021.4	98.4	1023	6 AB870246	AB870246 DNA encod
9	1020	98.3	1020	13 AB242578	AB242578 Human tra
10	1020	98.3	1020	13 AD873690	AD873690 Human tra
11	1018.4	98.1	1020	6 ABT04871	ABT04871 Human G P
12	1018.4	98.1	1020	12 AD030050	AD030050 Human GPC
13	1015.4	97.8	1017	4 AAH26919	AAH26919 Human G P
14	869	83.7	1138	6 AAH51012	AAH51012 Human ncp
15	869	83.7	1138	6 AB870245	AB870245 DNA encod
16	866.4	83.5	933	10 AD012791	AD012791 Human ncp
17	864	66.9	962	4 AAH51000	AAH51000 Human ncp
18	864	66.9	962	6 AB870233	AB870233 DNA encod
19	665.6	64.1	1101	4 AAC90246	AAC90246 Rat SNOR
20	653	62.9	1031	4 AAC90276	AAC90276 Mouse SNO

21	642.4	61.9	996	12 ADO30340	ADO30340 Mouse GPC
22	637.4	61.4	999	6 AB158317	AB158317 Rat trace
23	573	55.2	999	4 AAC90245	AAC90245 Human SNO
24	302.6	29.2	1041	12 ADO29146	ADO29146 Mouse nov
25	295.6	28.5	2273	2 AAH41998	AAH41998 Human HNH
26	292.8	28.2	1032	6 AB242548	AB242548 Human G P
27	292.8	28.2	1032	12 ADO29928	ADO29928 Human GPC
28	283	27.3	1029	12 ADO30218	ADO30218 Mouse GPC
29	279.8	27.0	1098	6 AB081178	AB081178 Human SNO
30	279.8	27.0	1098	10 ADE81114	ADE81114 Human SNO
31	276.6	26.6	1041	8 AB259359	AB259359 Human GPC
32	269.2	25.9	921	2 ABT94894	ABT94894 Human G P
33	269.2	25.9	921	10 ABV74592	ABV74592 Human G-P
34	269.2	25.9	1635	10 ADF70509	ADF70509 Orphan re
35	266	25.6	915	8 AB259360	AB259360 Human GPC
36	266	25.6	921	8 AB242547	AB242547 Human G P
37	266	25.6	921	12 ADO05721	ADO05721 Human G P
38	266	25.6	921	12 ADO28769	ADO28769 Human nov
39	263.2	25.4	921	12 ADO28075	ADO28075 Mouse nov
40	234.2	22.6	1044	12 ADO29149	ADO29149 Mouse nov
41	233	22.4	831	8 AB259361	AB259361 Mouse GPC
42	231.2	22.3	1034	8 AB269479	AB269479 Human G-P
43	231.2	22.3	1035	6 ABQ74527	ABQ74527 Human G P
44	231.2	22.3	1035	12 ADO30051	ADO30051 Human GPC
45	231.2	22.3	1038	6 ABT04870	ABT04870 Human G P

ALIGNMENTS

RESULT 1
AAC90247
ID AAC90247 standard; DNA; 1038 BP.

14-MAR-2001 (first entry)

Human SNORF3 receptor DNA.

SNORF3; inflammation; arthritis; neurological disorder; infection;
bone disease; respiratory disorder; asthma; cancer; cardiovascular; ds.

Homo sapiens.

MO200073449-A1.

07-DEC-2000.

26-MAY-2000; 2000MO-US014654.

28-MAY-1999; 99US-00322257.

06-OCT-1999; 99US-00413433.

(SYNA-) SYNAPTIC PHARM CORP.

Borowsky BE, Oggozalek KL, Jones KA;

WPI; 2001-025252/03.

Nucleic acid encoding a mammalian (human, rat and mouse) SNORF3 receptor
which is useful for designing drugs for treating conditions such as a
chronic and acute inflammation, arthritis, neurological disorders and
microbial infections.

Disclosure; Fig 5; 227p; English.

The present invention relates to a mammalian SNORF3 receptor. SNORF3
antagonists and agonists are used to treat abnormalities brought about by
increased or decreased activity of the mammalian SNORF3 receptor. The
receptor is useful as a tool for designing drugs for treating conditions
such as a chronic and acute inflammation, arthritis, neurological
disorders, microbial infections, bone diseases, respiratory disorders

XX 12-SEP-2001; 2001WO-US028455.
PF
XX 12-SEP-2000; 2000US-00659519.
PR 09-JUL-2001; 2001US-0303967P.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Bunzow JR, Grandy DK, Sonders M;
XX
XX MPI; 2002-371971/40.
DR P-PSDB; ABB80694.
XX
XX A novel mammalian biogenic amine receptor, useful in assays to identify
PT therapeutic agents for treating e.g., asthma and shock.
XX
XX Claim 6; Fig 1; 86pp; English.
XX
XX The invention relates to novel mammalian (human and rat) biogenic amine
CC receptors. The mammalian biogenic receptors can be used to screen for
CC compounds that bind to it, or to screen for compounds that inhibit a
CC mammalian trace amine receptor. The compounds identified by the assay
CC methods can be used to reduce sympathomimetic effects of enhanced trace
CC amine dependent synaptic transmission in mammals so they can be used to
CC to treat the peripheral effects of a drug that binds to or affects the
CC binding to trace amine receptors e.g. hyperthermia, rapid heart rate,
CC high blood pressure, migraine, cardiac arrhythmia, seizure, coma and
CC diabetes or to treat pathological conditions associated with elevated
CC levels of trace amines e.g. schizophrenia, depression, hypertension,
CC shock, cardiac arrhythmias, asthma, migraine, psychosis, anaphylactic
CC reactions and iatrogenic conditions. They are also useful for treating
CC drug addiction. The present sequence represents a human trace amine
CC receptor encoding DNA
XX
SQ Sequence 1125 BP; 313 A; 212 C; 211 G; 389 T; 0 U; 0 Other;
Query Match 99.1%; Score 1028.4; DB 6; Length 1125;
Best Local Similarity 99.4%; Pred. No. 2.5e-270;
Matches 1092; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TCAGGATGATGCGCTTTTGGCCAAATATTAATTTCTGTGTGAAAAACAACCTGG 60
DB 15 TCAGGATGATGCGCTTTTGGCCAAATATTAATTTCTGTGTGAAAAACAACCTGG 74
QY 61 TCAGGATGATGCGCTTTTGGCCAAATATTAATTTCTGTGTGAAAAACAACCTGG 120
DB 75 TCAGGATGATGCGCTTTTGGCCAAATATTAATTTCTGTGTGAAAAACAACCTGG 134
QY 121 GTTGGCAATCTGATAGTTATTTCTATATCACTTCAAAACAATTCATACCCCAACA 180
DB 135 GTTGGCAATCTGATAGTTATTTCTATATCACTTCAAAACAATTCATACCCCAACA 194
QY 181 AATTGGCTCATTCATTCATGCGCACTGAGACTTTCTTGGGGGTGCTGTCATGCTT 240
DB 195 AATTGGCTCATTCATTCATGCGCACTGAGACTTTCTTGGGGGTGCTGTCATGCTT 254
QY 241 TACAGATGATGATGATGCTGAGCACTGTTGATTTGGAGAACTCTTCTGTAATTT 300
DB 255 TACAGATGATGATGATGCTGAGCACTGTTGATTTGGAGAACTCTTCTGTAATTT 314
QY 301 CACACAGACCGACGATTAATGCTGAGCTCAGCTTCATTTTCATTTGCTTTCATCTCC 360
DB 315 CACACAGACCGACGATTAATGCTGAGCTCAGCTTCATTTTCATTTGCTTTCATCTCC 374
QY 361 ATTGACCGCTACTATGCTGTGTGATCCATGAGATATTAAGCCCAAGATGATATCTTG 420
DB 375 ATTGACCGCTACTATGCTGTGTGATCCATGAGATATTAAGCCCAAGATGATATCTTG 434
QY 421 GTTATTTGTTGATGATCTTCAATTAATTAATGAGAGTCCCTGTTTGGATTTGGAATG 480
DB 435 GTTATTTGTTGATGATCTTCAATTAATTAATGAGAGTCCCTGTTTGGATTTGGAATG 494
QY 481 ATCTTTCTGAGAGCTTAACCTCAAGGCGCTGAAGAGATATATTACAAACATGTTCACTGC 540

DB 495 ATCTTTCTGAGAGCTTAACCTCAAGGCGCTGAAGAGATATATTACAAACATGTTCACTGC 554
QY 541 AGAGAGGTTGCTGCTGCTCTTTAGCAAAATATCGGGGTCTGACCTTTATGACTTCT 600
DB 555 AGAGAGGTTGCTGCTGCTCTTTAGCAAAATATCGGGGTCTGACCTTTATGACTTCT 614
QY 601 TTTTATATACCTGATCTATTTATGTTATGTTCTATTAACAGAAATATCTTATCGCTAAA 660
DB 615 TTTTATATACCTGATCTATTTATGTTATGTTCTATTAACAGAAATATCTTATCGCTAAA 674
QY 661 GAACAGGCAAGATTAATTAATGATGCGCAATCAATGAGCTTCAAAATTTGATGAAATGAAA 720
DB 675 GAACAGGCAAGATTAATTAATGATGCGCAATCAATGAGCTTCAAAATTTGATGAAATGAAA 734
QY 721 AATGGAATTTCCAAAGCAAGAAAGAAAGCAAGCTGGAAGACATTTGGGGATTTGATGGGA 780
DB 735 AATGGAATTTCCAAAGCAAGAAAGAAAGCAAGCTGGAAGACATTTGGGGATTTGATGGGA 794
QY 781 GTTTTCCTAATATGCTGTCGCTTTCTTTATCTGTACAGTCATGACCTTTTCTTAC 840
DB 795 GTTTTCCTAATATGCTGTCGCTTTCTTTATCTGTACAGTCATGACCTTTTCTTAC 854
QY 841 TACATTATTCACCTACTTGTGAATGATGTTGATTTGGCTTGTGAACTCTACA 900
DB 855 TCAATTTATTCACCTACTTGTGAATGATGATTTGGCTTGTGAACTCTACA 914
QY 901 TTTATTCAGAGGTTAATGACATTTTCTATCTGTTAGAAAGACATGAAGATGATG 960
DB 915 TTTATTCAGAGGTTAATGACATTTTCTATCTGTTAGAAAGACATGAAGATGATG 974
QY 961 CTGTTTGGTAAATTTTCCAAAGATTCATCAGAGTGTAAATTTTGGAAATTTGACT 1020
DB 975 CTGTTTGGTAAATTTTCCAAAGATTCATCAGAGTGTAAATTTTGGAAATTTGACT 1034
QY 1021 TCATAGAAATTAATTAATTT 1038
DB 1035 TCATAGAAATTAATTAATTT 1052
RESULT 7
AAH51013
ID AAH51013 standard; DNA; 1023 BP.
XX
AC AAH51013;
XX
DT 28-AUG-2001 (first entry)
XX
DE Human nGPR56 coding sequence #3.
XX
KW G protein-coupled receptor; nGPR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder/attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective; ds.
XX
OS Homo sapiens.
XX
PN MO200136473-A2.
XX
PD 25-MAY-2001.
XX
PF 16-NOV-2000; 2000WO-US031581.
XX
PR 16-NOV-1999; 99US-0165838P.
PR 17-NOV-1999; 99US-0166071P.
PR 19-NOV-1999; 99US-0166678P.
PR 28-DEC-1999; 99US-0173396P.
PR 22-FEB-2000; 2000US-0184129P.

PR 28-FEB-2000; 2000US-0185421P.
 PR 28-FEB-2000; 2000US-0185542P.
 PR 02-MAR-2000; 2000US-0185543P.
 PR 03-MAR-2000; 2000US-0185544P.
 PR 09-MAR-2000; 2000US-0185545P.
 PR 17-MAR-2000; 2000US-0185546P.
 PR 21-MAR-2000; 2000US-0185547P.
 PR 20-APR-2000; 2000US-0190800P.
 PR 02-MAY-2000; 2000US-0190801P.
 PR 08-MAY-2000; 2000US-0203111P.
 PR 25-MAY-2000; 2000US-0207094P.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 XI Vogel I G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J.
 PI Schellin KA, Kaytes PS, Bannigan CM, Rufe V, Sejlitz T, Huff RM;
 XX
 DR WPI; 2001-389826/41.
 DR P-PSDB; MAG80973.
 XX
 PT New G protein-coupled receptor (ngPCR-x) and its encoding polynucleotide
 XX useful for diagnosing and treating e.g. schizophrenia.
 PS
 XX Claim 4; Page 92; 261pp; English.

CC The present invention relates to novel G protein-coupled receptors
 CC (ngPCR-x), where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27, 28,
 CC 31-38, 40, 41, 53-60) and their coding sequences. The present sequence is
 CC the coding sequence for one such G protein-coupled receptor. GPCR's are
 CC also known as seven transmembrane receptors and function in signal
 CC transduction. The ngPCR coding sequences are useful for screening a
 CC human to diagnose a disorder affecting the brain or a genetic
 CC predisposition, specifically schizophrenia. ngPCR's are useful for
 CC identifying compounds useful for treating schizophrenia. Detection of
 CC ngPCR in a sample is useful as a diagnostic tool for diseases or
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
 CC CNS disorders, infectious such as HIV-1, metabolic and cardiovascular
 CC diseases, proliferative disorders and hormonal disorders. Modulators of
 CC ngPCR activity have the utility for treating neurological disorders,
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
 CC disorder/attention deficit disorder), and neuronal disorders such as
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
 CC Additional disorders including inflammatory conditions (e.g. Crohn's
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.
 CC inflammatory bowel disease

Sequence 1023 BP; 277 A; 188 C; 199 G; 359 T; 0 U; 0 Other;

Query Match 98.4%; Score 1021.4; DB 4; Length 1023;
 Best Local Similarity 99.9%; Pred. No. 1.9e-268;
 Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGAAATGATGCTCTTTGGCAATATTAATATTTCTGTGAAAAACAATGTCGA 63
 DB 1 GGAAATGATGCTCTTTGGCAATATTAATATTTCTGTGAAAAACAATGTCGA 60
 QY 64 AATGATGCTGCTCTCTGACGTTTAAATGCTCATATTTGACCACTGCTT 123
 DB 61 AATGATGCTGCTCTCTGACGTTTAAATGCTCATATTTGACCACTGCTT 120
 QY 124 GGCAATCTGATATGTTATGTTCTATATCACTTCAAAACAATCAATCCCAACAAT 183
 DB 121 GGCAATCTGATATGTTATGTTCTATATCACTTCAAAACAATCAATCCCAACAAT 180
 QY 184 TGGCTCATTTTTCATGACGCTGTCGATCTTTCTGCGGCTGTCGATGCTTAC 243
 DB 181 TGGCTCATTTTTCATGACGCTGTCGATCTTTCTGCGGCTGTCGATGCTTAC 240
 QY 244 AGTATGATGATCTGTCGACGCTGTTGATTTTGAAGAAGCTCTGTAATTTAC 303
 DB 241 AGTATGATGATCTGTCGACGCTGTTGATTTTGAAGAAGCTCTGTAATTTAC 300

QY 304 ACAAGCACCAGATTATGCTGAGCTCAGCCTCCATTTTCCATTGCTTTGATCCATT 363
 DB 301 ACAAGCACCAGATTATGCTGAGCTCAGCCTCCATTTTCCATTGCTTTGATCCATT 360
 QY 364 GACCGCTACTATGCTGTGTGATCACTGAGATATTAAGCCAGATGAAATCTTGGTT 423
 DB 361 GACCGCTACTATGCTGTGTGATCACTGAGATATTAAGCCAGATGAAATCTTGGTT 420
 QY 424 ATTTGATGATCTTCAATTAAGTGAAGTCCGCTGCTGTTTGGATTTGATGATC 483
 DB 421 ATTTGATGATCTTCAATTAAGTGAAGTCCGCTGCTGTTTGGATTTGATGATC 480
 QY 484 TTTCTGAGCTTAACCTTCAAAAGCGCTGAAGATATATTAACAACATGTTCACTGAGA 543
 DB 481 TTTCTGAGCTTAACCTTCAAAAGCGCTGAAGATATATTAACAACATGTTCACTGAGA 540
 QY 544 GGAGGTGCTGCTGCTTTTGAAGAAATATCTGGGGTACTGACCTTTATGATCTTTT 603
 DB 541 GGAGGTGCTGCTGCTTTTGAAGAAATATCTGGGGTACTGACCTTTATGATCTTTT 600
 QY 604 TATATACCTGATCTATATGTTATGTTGCTATTAACAATATATCTTATGCTTAAAGA 663
 DB 601 TATATACCTGATCTATATGTTATGTTGCTATTAACAATATATCTTATGCTTAAAGA 660
 QY 664 CAGCAAGATTAATTAAGTATGATCCATGACAGCTCCAAATGATGAAATGAAAT 723
 DB 661 CAGCAAGATTAATTAAGTATGATCCATGACAGCTCCAAATGATGAAATGAAAT 720
 QY 724 GGAATTTTCAACAAGAAAGAAAGAGCTGTGAAGCATTTGGGGATTTGATGAGGATT 783
 DB 721 GGAATTTTCAACAAGAAAGAGAGCTGTGAAGCATTTGGGGATTTGATGAGGATT 780
 QY 784 TTTCTAATATGCTGTGCTCTTTCTTATCTGATGACGTCATGACCTTTTCTTCACTAC 843
 DB 781 TTTCTAATATGCTGTGCTCTTTCTTATCTGATGACGTCATGACCTTTTCTTCACTAC 840
 QY 844 ATTAATCCACTTACTTGAATGATGATGTTATGTTGGCTATGAACTGATCAATT 903
 DB 841 ATTAATCCACTTACTTGAATGATGATGTTATGTTGGCTATGAACTGATCAATT 900
 QY 904 AATCAATGCTTTATGATTTTCTATCTGCTTTTGAAGAAAGCACTGAAGATGATCTG 963
 DB 901 AATCAATGCTTTATGATTTTCTATCTGCTTTTGAAGAAAGCACTGAAGATGATCTG 960
 QY 964 TTTGGTAAATTTTCCAAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1023
 DB 961 TTTGGTAAATTTTCCAAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 QY 1024 TAG 1026
 DB 1021 TAG 1023

RESULT 8
 ABS70246
 ID ABS70246 standard; DNA; 1023 BP.
 XX
 AC ABS70246;
 XX
 DT 26-NOV-2002 (first entry)
 XX
 DE DNA encoding human G protein-coupled receptor, ngPCR-56, #2.
 XX
 KW Human; gene; de; G protein-coupled receptor; GPCR; ngPCR; beGPCR;
 KW seven transmembrane receptor; 7TM; mental disorder; diagnosis;
 KW genetic predisposition; brain; immune response; gene therapy;
 KW anxiety disorder; depression; bipolar disorder; schizophrenia;
 KW Huntington's disease; dyskinesia; manic depression; stroke;
 KW Parkinson's disease; Alzheimer's disease; diabetes; inflammation; wound;
 KW tranquilizer.
 XX
 OS Homo sapiens.

KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.
XX
XX Homo sapiens.
XX
XX WO200261087-A2.
XX
XX 08-AUG-2002.
XX
XX 19-DEC-2001; 2001WO-US050107.
XX
XX 19-DEC-2000; 2000US-0257144P.
XX
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
XX Burner GC, Roush CL, Brown JP;
XX
XX WPI; 2003-046718/04.
XX
XX P-BSDB; ABP81732.
XX
XX
XX New isolated antigenic peptides e.g., for G protein-coupled receptors
XX (GPCR), useful for diagnosing and designing drugs for treating conditions
XX in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
XX autoimmune diseases.
XX
XX
XX Disclosure; Fig 1; 523pp; English.
XX
XX
XX The present invention describes antigenic peptides (I) comprising: (a)
XX any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
XX acids. Also described: (1) an assay for the detection of a particular G
XX protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
XX and (2) an isolated antibody having high specificity and high affinity or
XX avidity for a particular GPCR. (I) can be used as GPCR modulators and in
XX gene therapy. The antigenic peptides for GPCRs are useful in detecting an
XX antibody against a particular GPCR, and in the production of specific
XX antibodies. The peptides and antibodies are also useful for detecting the
XX presence or absence of corresponding GPCRs. The antigenic peptides for
XX GPCRs and antibodies are useful for diagnosing and designing drugs for
XX treating immune-related diseases, growth-related diseases, cell
XX regeneration-related diseases, immunological-related cell proliferative
XX diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
XX atherosclerosis, bacterial, fungal, protozoan or viral infections,
XX osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
XX inflammation, allergies, Crohn's disease, diabetes, graft versus host
XX disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
XX anxiety, depression, schizophrenia, dementia, mental retardation, memory
XX loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
XX hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
XX any other disorder in which GPCRs are involved. The antibodies may be
XX used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
XX GPCR proteins given in ABP81675 to ABP82018, which are used in the
XX exemplification of the present invention
XX
XX
XX Sequence 1020 BP; 275 A; 188 C; 198 G; 359 T; 0 U; 0 Other;
SQ
Query Match 98.3%; Score 1020; DB 8; Length 1020;
Best Local Similarity 100.0%; Pred. No. 47e-268;
Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 ATGATGCCCTTTGGCACAATATTAATATTTCTGTGAAAAACAACCTGCTCAAT 66
DB 1 ATGATGCCCTTTGGCACAATATTAATATTTCTGTGAAAAACAACCTGCTCAAT 60
QY 67 GATGTCGTCCTCCCTGACATTAATGTCATTAATTTGACACCACTGTTGGC 126
DB 61 GATGTCGTCCTCCCTGACATTAATGTCATTAATTTGACACCACTGTTGGC 120
QY 127 AATGTGATAGTTATGTTTCTATATCAACTTCAAAACAATCTTACCCCAAAATTTG 186
DB 121 AATCTGATAGTTATGTTTCTATATCAACTTCAAAACAATCTTACCCCAAAATTTG 180
QY 187 CTCATTTATTCAGGACCACTGTGACTTTCTTGTGGGGTGTCTGATGCTTACAGT 246

DB 181 CTCATTCATTCAGGACCACTGTGACTTTCTTGTGGGGTGTCTGATGCTTACAGT 240
QY 247 ATGTGAGATCTGTGAGCACTGTGATTTTGGAGAAGTCTGTGAAAATTCACCA 306
DB 241 ATGTGAGATCTGTGAGCACTGTGATTTTGGAGAAGTCTGTGAAAATTCACCA 300
QY 307 AGCACCGACATTATGCTGAGCTGAGCTTCATTTTCCATTGTCTTTCATCTCATTTGAC 366
DB 301 AGCACCGACATTATGCTGAGCTGAGCTTCATTTTCCATTGTCTTTCATCTCATTTGAC 360
QY 367 CGCTACTATGCTGTGATGATCACTGAGATTAATAGCAAGTAATATCTTGGTTAT 426
DB 361 CGCTACTATGCTGTGATGATCACTGAGATTAATAGCAAGTAATATCTTGGTTAT 420
QY 427 TGTGTGATGATCTTCATTAGTTGAGTGTCCCTGCTGTTTGGATTGGAATGATCTTT 486
DB 421 TGTGTGATGATCTTCATTAGTTGAGTGTCCCTGCTGTTTGGATTGGAATGATCTTT 480
QY 487 CTGAGAGTTAACTTCAAGGCGCTGAGAGATATTTACAAACATGTTCACTGACAGAGA 546
DB 481 CTGAGAGTTAACTTCAAGGCGCTGAGAGATATTTACAAACATGTTCACTGACAGAGA 540
QY 547 GGTGCTCTGCTCTTCTTAGCAAAATATCTGGGTAAGTCACTTATGATCTTCTTTAT 606
DB 541 GGTGCTCTGCTCTTCTTAGCAAAATATCTGGGTAAGTCACTTATGATCTTCTTTAT 600
QY 607 ATACTGATCTATTTATGTTATGTCCTATTACAGAAATATATCTTATGCTTAAAGACAG 666
DB 601 ATACTGATCTATTTATGTTATGTCCTATTACAGAAATATATCTTATGCTTAAAGACAG 660
QY 667 GCAAGATTAATTTAGTGTGATGCAATCAGAGCTCCAAATTTGATTTGGAATGGAATGGA 726
DB 661 GCAAGATTAATTTAGTGTGATGCAATCAGAGCTCCAAATTTGATTTGGAATGGAATGGA 720
QY 727 ATTTCACAAAGCAAGAAAGAAAGCTGTGAAGCATTTGGGATTTGATGGAGTTTTC 786
DB 721 ATTTCACAAAGCAAGAAAGAAAGCTGTGAAGCATTTGGGATTTGATGGAGTTTTC 780
QY 787 CTATATGCTGTGTCCTTTCTTATCTGTACAGTATGAGCCCTTTCTTCACTACAT 846
DB 781 CTATATGCTGTGTCCTTTCTTATCTGTACAGTATGAGCCCTTTCTTCACTACAT 840
QY 847 ATTTCACCTTCTTGTATGATGTTGATTTGGCTTGAATCTTACATTTAAT 906
DB 841 ATTTCACCTTCTTGTATGATGTTGATTTGGCTTGAATCTTACATTTAAT 900
QY 907 CCAATGTTTATGATCTTTTCTATCTTGGTTTGAAGAAAGCACTGAAGATGATGCTGTT 966
DB 901 CCAATGTTTATGATCTTTTCTATCTTGGTTTGAAGAAAGCACTGAAGATGATGCTGTT 960
QY 967 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATTTTGGAAATTTGAGTTGATCATAG 1026
DB 961 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATTTTGGAAATTTGAGTTGATCATAG 1020
RESULT 10
AD573690
ID AD573690 standard; DNA; 1020 BP.
XX
XX AD573690;
XX
XX 16-DEC-2004 (first entry)
XX
XX Human trace amine receptor 1 (TA1) polypeptide encoding DNA.
XX
XX TA1; GPCR; trace amine receptor 1; antianaemic; antiinflammatory;
XX cardiovascular; neuroprotective; respiratory; uropathic;
XX G protein coupled receptor; gene therapy; human; CSRL; gene; ds.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FH 1..1020
XX CDS

```
PT      /*tag= a
PT      /product= "TAl receptor"
PT      /note= "trace amine receptor 1"
XX
XX      WO2004083851-A2.
XX
XX      30-SEP-2004.
XX
XX      04-MAR-2004; 2004WO-EP002182.
XX
XX      18-MAR-2003; 2003EP-00006101.
XX
XX      (FARB ) BAYER HEALTHCARE AG.
XX
XX      Golz S, Brueggemeier U, Geerts A;
XX
XX      MPI: 2004-691179/67.
XX      P-PSDB; ADS73691.
XX      GENBANK; NM_138327.
XX
XX      Screening for therapeutic agents, useful for treating e.g.,
XX      cardiovascular and inflammatory disorders, comprises contacting a test
XX      compound with trace amine receptor 1 (TAl) polypeptide and detecting
XX      their binding.
XX
XX      Disclosure; SEQ ID NO 1; 117bp; English.
XX
XX      The invention relates to screening for therapeutic agents and involves
XX      contacting a test compound with trace amine receptor 1 (TAl) polypeptide
XX      or polynucleotide and detecting binding of the test compound to TAl
XX      polypeptide or polynucleotide, or determining TAl polypeptide activity at
XX      a certain test compound concentration or in the absence of the test
XX      compound and at a different concentration of the test compound. The
XX      therapeutic agents are useful in treating disease such as cardiovascular
XX      disorders, inflammatory diseases, hematological disorders, respiratory
XX      diseases, neurological disorders or urological disorders in a mammal. The
XX      regulators of TAl are useful in the preparation of a pharmaceutical
XX      composition for treating the above diseases and also useful for the
XX      regulation of TAl activity in a mammal having the disease. The nucleotide
XX      sequences encoding TAl are useful as hybridization probes, in
XX      constructing oligomers for PCR, for chromosome and gene mapping, in the
XX      recombinant production of TAl, in generating antisense DNA or RNA and in
XX      molecular biology techniques that have not yet been developed. TAl are
XX      useful for immunizing a mammal to produce polyclonal antibodies and for
XX      diagnostic purposes. The present sequence represents a DNA encoding a
XX      human TAl receptor, a seven transmembrane G protein coupled receptor
XX      polypeptide.
XX
XX      Sequence 1020 BP; 275 A; 188 C; 198 G; 359 T; 0 U; 0 Other;
XX
XX      Query Match      98.3%; Score 1020; DB 13; Length 1020;
XX      Best Local Similarity 100.0%; Pred. No. 4.7e-268;
XX      Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      7 ATGATGCCCTTTGGCACAATATTAATTAATTTCTGTGTGAAAAAACAATGTCGAAT 66
XX      1 ATGATGCCCTTTGGCACAATATTAATTAATTTCTGTGTGAAAAAACAATGTCGAAT 60
XX
XX      67 GATGTCGCTGCTTCCCTGTACAGTTAATGTCATATATTGTGACCACTCGTTGGC 126
XX      61 GATGTCGCTGCTTCCCTGTACAGTTAATGTCATATATTGTGACCACTCGTTGGC 120
XX
XX      127 AATCTGATAGTATTGTTCTATATCAGACTTCAACCACTCATGCCCAACAATTGG 186
XX      121 AATCTGATAGTATTGTTCTATATCAGACTTCAACCACTCATGCCCAACAATTGG 180
XX
XX      187 CTCATTCAATTCATGGCCACTGTGACCTTCTTCTGGGGTGTCTGTCATGCTTACAGT 246
XX      181 CTCATTCAATTCATGGCCACTGTGACCTTCTTCTGGGGTGTCTGTCATGCTTACAGT 240
XX
XX      247 ATGGTGAGATCTGCTGAGCAGCTGTTGTATTTTGGAGAAAGTCTTCTGTAATTCACACA 306
XX      241 ATGGTGAGATCTGCTGAGCAGCTGTTGTATTTTGGAGAAAGTCTTCTGTAATTCACACA 300
XX
XX      Db
```

```
QY      307 AGCACCAGCATTTATGCTGAGCTCAGCCTCCATTTTTCATTTGTCTTTCATCTCATTTGAC 366
XX      |||
XX      301 AGCACCAGCATTTATGCTGAGCTCAGCCTCCATTTTTCATTTGTCTTTCATCTCATTTGAC 360
XX      |||
XX      QY      367 CGCTACTATGCTGTGTGTGATTCAGTGAATTTAAAGCCAGATGAATATCTTGGTATT 426
XX      |||
XX      Db      361 CGCTACTATGCTGTGTGTGATTCAGTGAATTTAAAGCCAGATGAATATCTTGGTATT 420
XX      |||
XX      QY      427 TGTGTGATGATCTTATAGTTAGTGGATGCCCTGCTGTTTGGATTTGGAATGATCTTT 486
XX      |||
XX      Db      421 TGTGTGATGATCTTATAGTTAGTGGATGCCCTGCTGTTTGGATTTGGAATGATCTTT 480
XX      |||
XX      QY      487 CTGAGACTTAACTTCAAGGCGCTGAGAGATATATTCAAAATGTTCACTGCAGAGAGA 546
XX      |||
XX      Db      481 CTGAGACTTAACTTCAAGGCGCTGAGAGATATATTCAAAATGTTCACTGCAGAGAGA 540
XX      |||
XX      QY      547 GGTTCCTCTGCTCTTTTACGAAATATCTGGGGTACTGACCTTTATGACTTCTTTTAT 606
XX      |||
XX      Db      541 GGTTCCTCTGCTCTTTTACGAAATATCTGGGGTACTGACCTTTATGACTTCTTTTAT 600
XX      |||
XX      QY      607 ATACCTGATCTATATGTTATGTTGTCTATTACAGATATATCTTATGCTTAAAGACAG 666
XX      |||
XX      Db      601 ATACCTGATCTATATGTTATGTTGTCTATTACAGATATATCTTATGCTTAAAGACAG 660
XX      |||
XX      QY      667 GCAGATTTAATTAGTATGATGCCAATCAGAAAGCTCCAAATTGGAATTGGAATGAAATGGA 726
XX      |||
XX      Db      661 GCAGATTTAATTAGTATGATGCCAATCAGAAAGCTCCAAATTGGAATTGGAATGAAATGGA 720
XX      |||
XX      QY      727 ATTTCAAGAAAGCAAGAAAGAAAGAGCTGTGAAGACATTTGGGATTTGTATGGAGATTTC 786
XX      |||
XX      Db      721 ATTTCAAGAAAGCAAGAAAGAAAGAGCTGTGAAGACATTTGGGATTTGTATGGAGATTTC 780
XX      |||
XX      QY      787 CTAATATGCTGGTGCCTTTCTTATCTGACATGATGAGACCTTTCTTCTACATCAT 846
XX      |||
XX      Db      781 CTAATATGCTGGTGCCTTTCTTATCTGACATGATGAGACCTTTCTTCTACATCAT 840
XX      |||
XX      QY      847 ATTCACCTACTTTGAATGATGTGTGATTTGTTGGCTACTTGAACCTTACATTTTAT 906
XX      |||
XX      Db      841 ATTCACCTACTTTGAATGATGTGTGATTTGTTGGCTACTTGAACCTTACATTTTAT 900
XX      |||
XX      QY      907 CCAATGTTTATGATATTTTCTATCTGTTTGTAGAAAGCAGTGAAGATGCTGTTT 966
XX      |||
XX      Db      901 CCAATGTTTATGATATTTTCTATCTGTTTGTAGAAAGCAGTGAAGATGCTGTTT 960
XX      |||
XX      QY      967 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATATTTTGGAAATGATGTCATAG 1026
XX      |||
XX      Db      961 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATATTTTGGAAATGATGTCATAG 1020
XX      |||
XX
XX      RESULT 11
XX      ABT04871
XX      ID      ABT04871 standard; cDNA; 1020 BP.
XX
XX      ABT04871;
XX
XX      AC      ABT04871;
XX      XX
XX      DT      11-OCT-2002 (first entry)
XX      XX
XX      DE      Human G protein coupled receptor hRUP33 coding sequence.
XX      XX
XX      KW      Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
XX      KW      hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37; gene; ss.
XX      OS      Homo sapiens.
XX      PN      WO200242461-A2.
XX      PD      30-MAY-2002.
XX      PF      26-NOV-2001; 2001WO-US044386.
XX      PR      27-NOV-2000; 2000US-0253404P.
XX      PR      12-DEC-2000; 2000US-0255366P.
```

Query Match 98.1%; Score 1018.4; DB 6; Length 1020;
 Best Local Similarity 99.9%; Pred. No. 1.3e-267;
 Matches 1019; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy		607	ATACCTGGATTCTTATGTTATGTGCCTATTCAGAAATATATTTTATCGCTTAAGAACAAG	666		
Dz		601	ATRCCCTGGATCTTATTATGTTATGTCTATTCAGAAATATATCTTATCGCTTAAGAACAAG	660		
Qy		667	GCAAGATTATATTAGTGATGCCAATCAGAAAGCTCCAAATTGGAATTGAAAATGGA	726		
Dz		661	GCAGAATTATATTAGTGATGCCAATCAGAAAGCTCCAAATTGGAATTGAAAATGGA	720		
Qy		727	ATTTCACAAAGAAAAAGAAAGAACCTGGAGACTTTGGGATTTGTATGGAGTTTTTC	786		
Dz		721	ATTTCACAAAGAAAAAGAAAGAACCTGGAGACTTTGGGATTTGTATGGAGTTTTTC	780		
Qy		787	CTAATATGTGTGTCCTTTCTTTATCTGTACAGTCATGGACCCTTTTCTTACTACAT	846		
Dz		781	CTAATATGTGTGTCCTTTCTTTATCTGTACAGTCATGGACCCTTTTCTTACTACAT	840		
Qy		847	ATTCACCTACTTTGAATGATGTGTGATTGGTTGGCTACTTGAACCTTACATTAT	906		
Dz		841	ATTCACCTACTTTGAATGATGTGTGATTGGTTGGCTACTTGAACCTTACATTAT	900		
Qy		907	CCAATGTTTATGCAATTTTCTATCCTTGTTTGAAGAAAGCATGAATGATGCTGTT	966		
Dz		901	CCAATGTTTATGCAATTTTCTATCCTTGTTTGAAGAAAGCATGAATGATGCTGTT	960		
Qy		967	GGTAAAAATTTTCCAAAAAGATTCATCCAGGTGTAATTTTGGAAATGAGTTCATG	1026		
Dz		961	GGTAAAAATTTTCCAAAAAGATTCATCCAGGTGTAATTTTGGAAATGAGTTCATG	1020		
<hr/>						
RESULT 12						
ID	AD030050	standard; cDNA; 1020 BP.				
AC	AD030050;					
XX	29-JUL-2004	(first entry)				
DE	Human GPCR TARI polynucleotide, SEQ ID NO:1152.					
KW	G protein-coupled receptor; GPCR; drug screening; diagnosis;					
KM	transgenic mouse; neurological disorder; adrenal gland disorder;					
KW	colon disorder; intestinal disorder; cardiovascular disorder;					
KM	muscular disorder; blood disorder; immune disorder; bone disorder;					
KW	joint disorder; metabolic disorder; nutritive disorder; cancer;					
KM	kidney disorder; liver disorder; lung disorder; breast disorder;					
KW	ovary disorder; uterus disorder; prostate disorder; testis disorder;					
KM	skin disorder; stomach disorder; pancreas disorder; spleen disorder;					
KW	tlymus disorder; thyroid disorder; vasotropic; antiangiinal; antiahythmic;					
KM	Cytostatic; antinflammatorcy; vasocropic; antianginal; antidiabetic;					
KW	CNS; central nervous system; respiratory; antidiarhoeic; antidiabetic;					
KM	viruside; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;					
KW	dermatological; antihijer; antithyroid; antiallergic; anorectic;					
KM	immunopressiveive; nephrotropic; gene therapy; GPCR modulator; human;					
KW	gene; ss.					
OS	Homo sapiens.					
XX	WO2004040000-AZ.					
PD	13-MAY-2004.					
Pf	09-SEP-2003; 2003WO-US028226.					
PR	09-SEP-2002; 2002US-0409303P.					
PA	09-APR-2003; 2003US-0461329P.					
PI	(PRIM-) PRIMAL INC.					
PI	Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F,					
PI	Madisen L, Mollwain KL, Pavlova MN, Vassiliadis D, Zeng H;					
XX	WPI, 2004-390329/36.					

DR P-PsDB; ADO29662.
XX Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
XX Claim 151; SEQ ID NO 1152; 542bp; English.
XX
CC The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridize to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification; those sequences
CC not shown were obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences.
XX
XX Sequence 1020 BP; 276 A; 188 C; 197 G; 359 T; 0 U; 0 Other;
SQ

Query Match 98.1%; Score 1018.4; DB 12; Length 1020;
Best Local Similarity 99.9%; Pred. No. 1.3e-267;
Matches 1019; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ATGATGCCCTTTGGCACAATATTAATTAATTTCTGTGTGAAAAACAACAGTGTCAAT 66
XX
XX 1 ATGATGCCCTTTGGCACAATATTAATTAATTTCTGTGTGAAAAACAACAGTGTCAAT 60
DB

QY 67 GATGCCCTTTGGCACAATTAATGTCATTAATTTGACCACTGCTGGC 126
XX
DB 61 GATGCCCTTTGGCACAATTAATGTCATTAATTTGACCACTGCTGGC 120

QY 127 AATCGATAGTATTTGTTCTATATCACTCAAAACAATTCAACCCCAAAATGG 186
XX
DB 121 AATCGATAGTATTTGTTCTATATCACTCAAAACAATTCAACCCCAAAATGG 180

QY 187 CTCATTCATTCATGAGCACTGTGACCTTTCTTGCGGGTGTCTGTCATGCTTACAGT 246
XX
DB 181 CTCATTCATTCATGAGCACTGTGACCTTTCTTGCGGGTGTCTGTCATGCTTACAGT 240

QY 247 ATGTCGAGATCTGTCGAGACACTGTTGTAATTTTGGAGAGCTTCTGTAATTTACACA 306
XX
DB 241 ATGTCGAGATCTGTCGAGACACTGTTGTAATTTTGGAGAGCTTCTGTAATTTACACA 300

QY 307 AGCAGCGACATTAATGTCGAGCTCAGGCTCATTTTCCATTTGCTTTCATTCACATTGAC 366
XX
DB 301 AGCAGCGACATTAATGTCGAGCTCAGGCTCATTTTCCATTTGCTTTCATTCACATTGAC 360

QY 367 CGCTACTATGCTGTGTGTGATCCACTGAGATATAAGCCAAAGATGAATATCTTGTTAT 426
XX
XX

DB 361 CGTACTATGCTGTGTGTGATCCACTGAGATATAAGCCAAAGATGAATATCTTGTTAT 420
QY 427 TGTGTGATGATCTTCATTAATGAGAGTCCCTGCTGTTTGGATTTGGAATGATCTT 486
XX
DB 421 TGTGTGATGATCTTCATTAATGAGAGTCCCTGCTGTTTGGATTTGGAATGATCTT 480

QY 487 CTGAGCTAAACCTTCAAGGCGCTGAGAGATATATTAACAACATGTTCACTGCAGAGA 546
XX
DB 481 CTGAGCTAAACCTTCAAGGCGCTGAGAGATATATTAACAACATGTTCACTGCAGAGA 540

QY 547 GATGCTGTGCTCTTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTTTAT 606
XX
DB 541 GATGCTGTGCTCTTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTTTAT 600

QY 607 ATACTGATCTATTAATGTTATGTTCTATTAACAGATATATCTTATGCTTAAAGACG 666
XX
DB 601 ATACTGATCTATTAATGTTATGTTCTATTAACAGATATATCTTATGCTTAAAGACG 660

QY 667 GCAAGATTAATTAATGATGCAAGAGCTCCAAATTTGATTTGGAATGAAATGGA 726
XX
DB 661 GCAAGATTAATTAATGATGCAAGAGCTCCAAATTTGATTTGGAATGAAATGGA 720

QY 727 ATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTATGAGAGTTTC 786
XX
DB 721 ATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGATTTGTATGAGAGTTTC 780

QY 787 CTAATATGCTGTGCTCTTTTATCTGTAAGCAAGACCTTTCTTACATCAAT 846
XX
DB 781 CTAATATGCTGTGCTCTTTTATCTGTAAGCAAGACCTTTCTTACATCAAT 840

QY 847 ATTCACCTACTTTGAATGATGTGTGATTTGGTTGCTACTTGAACCTACATTTAT 906
XX
DB 841 ATTCACCTACTTTGAATGATGTGTGATTTGGTTGCTACTTGAACCTACATTTAT 900

QY 907 CCAATGTTTATGCAATTTTCTATCTGTTTGTAGAAAAGCACTGAAGATGCTGTT 966
XX
DB 901 CCAATGTTTATGCAATTTTCTATCTGTTTGTAGAAAAGCACTGAAGATGCTGTT 960

QY 967 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATATTTTGGAAATGAGTCAATG 1026
XX
DB 961 GGTAAATTTTCCAAAAGATTCATCCAGGTGAATATTTTGGAAATGAGTCAATG 1020

RESULT 13
AAH26919
ID AAH26919 standard; cDNA, 1017 BP.
XX
XX AAH26919;
XX
XX 21-DEC-2001 (first entry)
XX
DB Human G protein coupled receptor cDNA.
XX
XX G protein coupled receptor; GPCR; aminergic receptor;
XX signal transduction; gene therapy; human; ss.
XX
XX Homo sapiens.
XX
XX WO200172841-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US009660.
XX
XX 27-MAR-2000; 2000US-0192311P.
XX
XX 04-AUG-2000; 2000US-00633145.
XX
XX (PEKE) PE CORP NY.
XX
XX Kodira C, Cravchik A, Di Francesco V, Beasley EM;
XX
XX WPI, 2001-611617/70.
XX
XX P-PsDB; AAB82970.
DR

XX New human G-protein coupled receptor, useful for identifying specific
PT modulators, potential therapeutic agents, is related to the aminergic
PT receptor family.

PS Claim 4; Fig 1; 60pp; English.

The present sequence is that of cDNA encoding novel human G protein coupled receptor (GPCR, see A4882970) that is related to aminergic receptor family. GPCRs, particularly members of the aminergic receptor family, are a major target for drug action and development. The nucleotide sequences and polypeptides of the invention are therefore useful for the development of human therapeutic targets, can aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents. The present GPCR is expressed in foetal brain, brain, placenta, liver, stomach and kidney and is involved in signal transduction pathways. By targeting an agent to modulate the GPCR, the signalling activity and biological process mediated by the receptor can be agonised or antagonised in specific cells and tissues in order to modulate its biological activity in a therapeutic context (mammalian therapy) or toxic context (anti-cell therapy e.g. anti-cancer therapy). Nucleic acids encoding the GPCR are useful for recombinant expression of the protein, and as a source of primers and probes (for diagnosis), antisense sequences and ribozymes (for therapy). They can also be used for identifying modulators of expression, monitoring gene expression during therapy, identifying mutations in the encoding gene, construction of gene therapy vectors, and preparing transgenic animals, used to study GPCR function and to identify modulators

Sequence 1017 BP; 275 A; 188 C; 196 G; 358 T; 0 U; 0 Other;

Query Match	97.8%;	Score 1015.4;	DB 4;	Length 1017;
Best Local Similarity	99.9%;	Pred. No. 8.4e-267;		
Matches 1016: Conservative	0;	Mismatches	1	T-2-1

QY 10 AAGCCCTTTGGCACAATAATTAATTAATTTCTGTGTGAAAACAACGTGTCAATGAT 69
 Db 1 AAGCCCTTTGGCACAATAATTAATTAATTTCTGTGTGAAAACAACGTGTCAATGAT 60
 QY 70 GTCCGTGCTTCCCTGTACAGTTAATGTGCTCATTAATCTGACCACACTGTGTGGCAAT 129
 Db 61 GTCCGTGCTTCCCTGTACAGTTAATGTGCTCATTAATCTGACCACACTGTGTGGCAAT 120
 QY 130 CTGATAGTATGTGTTCTTAATATCACACTTCAACAACTTCATACCCCAACAAATTTGCTC 189
 Db 121 CTGATAGTATGTGTTCTTAATATCACACTTCAACAACTTCATACCCCAACAAATTTGCTC 180
 QY 190 ATTCAATTCATGCGCACCTGTGACCTTTCTTCTGGGGGTGCTGGGTGATGCTTACAGTATG 249
 Db 181 ATTCAATTCATGCGCACCTGTGACCTTTCTTCTGGGGGTGCTGTGTATGCTTACAGTATG 240
 QY 250 GTGAAATGTGCTGAGACACTGTTGGTAATTTTGGAGAAGTCTTGTGAAATTCACACAAGC 309
 Db 241 GTGAAATGTGCTGAGACACTGTTGGTAATTTTGGAGAAGTCTTGTGAAATTCACACAAGC 300
 QY 310 ACCGACATTATGCTGAGCTGACGCCCTCCATTTTCCATTTGTGCTTCATCTCATATGACCGC 369
 Db 301 ACCGACATTATGCTGAGCTGACGCCCTCCATTTTCCATTTGTGCTTCATCTCATATGACCGC 360
 QY 370 TACTATGTGTGTGTGATCCACTGAGATATAAAGCCAAGATGAATATCTTGGTATTTGT 429
 Db 361 TACTATGTGTGTGTGATCCACTGAGATATAAAGCCAAGATGAATATCTTGGTATTTGT 420
 QY 430 GTGATGATCTTCATTAATGTTGAGGTGTCCTGCTGTTTTTGGATTTGGAAATATCTTTCTG 489
 Db 421 GTGATGATCTTCATTAATGTTGAGGTGTCCTGCTGTTTTTGGATTTGGAAATATCTTTCTG 480
 QY 490 GAGCTAAACTTCAAAGGCGCTGAAAGAGATATATTACAAACATGTTCACTGACAGAGAGGT 549
 Db 481 GAGCTAAACTTCAAAGGCGCTGAAAGAGATATATTACAAACATGTTCACTGACAGAGAGGT 540
 QY 550 TGCTCTGCTCTTTAGCAAAATATCTGGGGTACTGACCTTATGACTCTTTTATATATA 609

PR 08-MAY-2000; 2000US-020311JP.
PR 25-MAY-2000; 2000US-0207094P.
XX
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Vogel I G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J,
PI Schellin KA, Kayles PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM,
XX
XX WPI: 2001-389826/41.
DR P-PSDB; ABG80972.
XX
XX
PT New G protein-coupled receptor (ngPCR-x) and its encoding polynucleotide
PT useful for diagnosing and treating e.g. schizophrenia.
XX
XX
PS Claim 4; Page 91; 261pp; English.
XX
XX The present invention relates to novel G protein-coupled receptors
CC (ngPCR); where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27, 28,
CC 31-38, 40, 41, 53-60) and their coding sequences. The present sequence is
CC the coding sequence for one such G protein-coupled receptor. GPCRs are
CC also known as seven transmembrane receptors and function in signal
CC transduction. The ngPCR coding sequences are useful for screening a
CC human to diagnose a disorder affecting the brain or a genetic
CC predisposition, specifically schizophrenia. ngPCR are useful for
CC identifying compounds useful for treating schizophrenia. Detection of
CC ngPCR in a sample is useful as a diagnostic tool for diseases or
CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular
CC diseases, proliferative disorders and hormonal disorders. Modulators of
CC ngPCR activity have the utility for treating neurological disorders,
CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
CC disorder/attention deficit disorder), and neuronal disorders such as
CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
CC Additional disorders include inflammatory conditions (e.g. Crohn's
CC disease), rheumatoid arthritis, autoimmune disorders, cancers, e.g.
CC respiratory ailments such as asthma, and inflammatory diseases e.g.
CC inflammatory bowel disease
XX
XX
XX Sequence 1138 BP; 339 A; 207 C; 220 G; 372 T; 0 U; 0 Other;
XX
XX
XX Query Match 83.7%; Score 869; DB 4; Length 1138;
XX Bees Local Similarity 100.0%; Pred. No. 8.6e-227;
XX Matches 869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
QY 1 TCAGGAATGATGCTTCTTCCCAACAATATTAATTAATTTCTGCTGTAAGAAACAACGG 60
DB 269 TCAGGAATGATGCTTCTTCCCAACAATATTAATTAATTTCTGCTGTAAGAAACAACGG 328
QY 61 TCAGGAATGATGCTTCTTCCCAACAATATTAATTAATTTCTGCTGTAAGAAACAACGG 120
DB 329 TCAGGAATGATGCTTCTTCCCAACAATATTAATTAATTTCTGCTGTAAGAAACAACGG 388
QY 121 GTTGGCAATCTGATAGTTATTTCTTATATCAACTTCAACAACCTTCAATCCCAACA 180
DB 389 GTTGGCAATCTGATAGTTATTTCTTATATCAACTTCAACAACCTTCAATCCCAACA 448
QY 181 AATTGGCTCATTCATTCATGACCTGACCTTCTTCTGCGGGGCTGTCATGCTT 240
DB 449 AATTGGCTCATTCATTCATGACCTGACCTTCTTCTGCGGGGCTGTCATGCTT 508
QY 241 TACAGATGATGATGCTGCTGACACTGTTGGTATTTGGAGAGTCTTCTGTAATTT 300
DB 509 TACAGATGATGATGCTGCTGACACTGTTGGTATTTGGAGAGTCTTCTGTAATTT 568
QY 301 CACACAGCAGCAGCATTTATGCTGAGCTCAGCTTCATTTTCCATTTGCTTTCATCTCC 360
DB 569 CACACAGCAGCAGCATTTATGCTGAGCTCAGCTTCATTTTCCATTTGCTTTCATCTCC 628
QY 361 ATTGACCGCTACTATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 420
DB 629 ATTGACCGCTACTATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 688
QY 421 GTTATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

DB 689 GTTATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
QY 481 ATCTTTCGAGCTTAAATTTCAAGCGCTGAAGAGATATTTCAACATGTTCACTGC 540
DB 749 ATCTTTCGAGCTTAAATTTCAAGCGCTGAAGAGATATTTCAACATGTTCACTGC 808
QY 541 AGAGAGGTTGCTGCTGCTTCTTGAACAATATCTGGGCTACTGACCTTTATGACTTCT 600
DB 809 AGAGAGGTTGCTGCTGCTTCTTGAACAATATCTGGGCTACTGACCTTTATGACTTCT 868
QY 601 TTTTATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 869 TTTTATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 928
QY 661 GAACGAGCAATATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 929 GAACGAGCAATATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 988
QY 721 AATGGAATTTTCAACAAGCAAGAAAGAAAGCTGGAAGATGATGATGATGATGATGATGAT 780
DB 989 AATGGAATTTTCAACAAGCAAGAAAGAAAGCTGGAAGATGATGATGATGATGATGATGAT 1048
QY 781 GTTTCCTTAATATGCTGCTGCTTCTTATCTGATGATGATGATGATGATGATGATGAT 840
DB 1049 GTTTCCTTAATATGCTGCTGCTTCTTATCTGATGATGATGATGATGATGATGATGAT 1108
QY 841 TACATTAATTCACCACTTGAATGATGAT 869
DB 1109 TACATTAATTCACCACTTGAATGATGAT 1137
XX
XX
XX RESULT 15
XX ABS70245
XX ID ABS70245 standard; DNA; 1138 BP.
XX
XX ABS70245;
XX
XX
XX 26-NOV-2002 (first entry)
XX
XX
XX DNA encoding human G protein-coupled receptor, ngPCR-56, #1.
XX
XX DE Human; gene; ds; G protein-coupled receptor; GPCR; ngPCR; beGPCR;
XX KW ng protein coupled receptor; communication; serpentine structure;
XX KW seven transmembrane receptor; 7TM; mental disorder; diagnosis;
XX KW genetic predisposition; brain; immune response; gene therapy;
XX KW anxiety disorder; depression; bipolar disorder; schizophrenia;
XX KW Huntington's disease; dyskinnesia; manic depression; stroke;
XX KW Parkinson's disease; Alzheimer's disease; diabetes; inflammation; wound;
XX KW tranquiliser.
XX
XX OS Homo sapiens.
XX
XX XX
XX PN WO200264789-A1.
XX
XX XX
XX PD 22-AUG-2002.
XX
XX PF 14-FEB-2001; 2001WO-US004641.
XX
XX PR 14-FEB-2001; 2001WO-US004641.
XX
XX PA (PHAA) PHARMACIA & UPJOHN CO.
XX
XX PI Lind P, Parodi LA, Vogel I G, Wood LS;
XX
XX WPI: 2002-674879/72.
XX P-PSDB; ABG93790.
XX
XX
XX New nucleic acids and polypeptides of the ng protein-coupled receptor,
XX useful for treating or diagnosing a mental disorder or a disorder
XX affecting the brain, e.g. anxiety disorders, schizophrenia, stroke or
XX Parkinson's disease.
XX

PS Example 1; Page 86; 244pp; English.

XX The invention discloses an isolated human polypeptide, and encoding
CC nucleic acid, for a G protein-coupled receptor (GPCR), particularly the
CC ng protein coupled receptor-14 (NGPCR-14). GPCRs are vital in the
CC communication between cells and their environment and are characterised
CC by a serpentine structure that passes through the cell membrane seven
CC times, hence the reason such receptors are sometimes called seven
CC transmembrane receptors (7TM). The polynucleotides and polypeptides are
CC useful for identifying an NGPCR allelic variant that correlates with a
CC mental disorder, for isolating an antibody that binds to an epitope of
CC the polypeptide, for identifying a compound that binds to an epitope of
CC polynucleotide and/or mediates its biological activity, for screening a
CC human subject to diagnose a disorder, or a genetic predisposition to a
CC disorder, affecting the brain or a genetic disposition to the disorder,
CC and for identifying compounds useful for the treatment of a mental disorder
CC and for identifying a compound useful as a modulator of binding between
CC NGPCR-14 and a binding partner of NGPCR-14. The polypeptide is also
CC useful for inducing an immune response in a mammal. The nucleic acid or
CC polypeptide is particularly useful, using gene therapy, for treating e.g.
CC anxiety disorders, depression, bipolar disorder, schizophrenia,
CC Huntington's disease, dyskinesias, manic depression, stroke, Parkinson's
CC disease or Alzheimer's disease. The nucleic acid and polypeptide may also
CC be used for treating diabetes, inflammation or wounds. The sequences
CC presented in ABS70202-ABS70248, ABS70338 and ABS70243 are the DNAs
CC encoding the NGPCRs (also referred to as beGPCRs)

SQ Sequence 1138 BP; 339 A; 207 C; 220 G; 372 T; 0 U; 0 Other;

Query Match 83.7%; Score 869; DB 6; Length 1138;
Best Local Similarity 100.0%; Pred. No. 8.6e-227;

Matches 869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TCAGGATATATGCTCTTTTCCACATATATATATTTCTGTGTGAAAAACAACACTGG 60
DB 269 TCAGGATATATGCTCTTTTCCACATATATATATTTCTGTGTGAAAAACAACACTGG 328
QY 61 TCAGGATATGTCGTCCTTCCCTGACATTTAATGTCCTATATTTGACCACTC 120
DB 329 TCAGGATATGTCGTCCTTCCCTGACATTTAATGTCCTATATTTGACCACTC 388
QY 121 GTTGCAATCTGATATGTTATTTCTATATACACTTCAACCACTTCATACCCACACA 180
DB 389 GTTGCAATCTGATATGTTATTTCTATATACACTTCAACCACTTCATACCCACACA 448
QY 181 AATTGGCTCATTCATTCATGCGCACTGTGACCTTCTTGTGGGTCGTGTCATGCT 240
DB 449 AATTGGCTCATTCATTCATGCGCACTGTGACCTTCTTGTGGGTCGTGTCATGCT 508
QY 241 TACAGTATGATGATCTGCTGAGCACTGTTGATTTTGGAGAACTTCTGTTAAAT 300
DB 509 TACAGTATGATGATCTGCTGAGCACTGTTGATTTTGGAGAACTTCTGTTAAAT 568
QY 301 CACACAGACACGACATTAATGCTGAGCTCAGCTCCATTTTCCATTTGTCTTCATCTCC 360
DB 569 CACACAGACACGACATTAATGCTGAGCTCAGCTCCATTTTCCATTTGTCTTCATCTCC 628
QY 361 ATTGACCGCTACTATGCTGTGTGATCTGATGATTAAGCCAAAGATGATATCTTGG 420
DB 629 ATTGACCGCTACTATGCTGTGTGATCTGATGATTAAGCCAAAGATGATATCTTGG 688
QY 421 GTTATTTGTGATGATCTTCATTAATGATGAGTGCCCTGCTGTTTGTGATTTGGAATG 480
DB 689 GTTATTTGTGATGATCTTCATTAATGATGAGTGCCCTGCTGTTTGTGATTTGGAATG 748
QY 481 ATCTTTTGTGAGCTTAACTTCAAAAGCGCTGAGAGATATTAACAAACATGTTCACTGC 540
DB 749 ATCTTTTGTGAGCTTAACTTCAAAAGCGCTGAGAGATATTAACAAACATGTTCACTGC 808
QY 541 AGAGAGAGTTGCTGCTGCTTTTGTAGCAAAATATCTGGGGTACCTGACCTTTAGACTTCT 600
DB 809 AGAGAGAGTTGCTGCTGCTTTTGTAGCAAAATATCTGGGGTACCTTATGACTTCT 868
```

```
QY 601 TTTTATATACCTGATCTTATATGTTATGTTCTATTAACAGATATATCTTATGCTTAAA 660
DB 869 TTTTATATACCTGATCTTATATGTTATGTTCTATTAACAGATATATCTTATGCTTAAA 928
QY 661 GAACAGCAAGATTAATTAATGATGATGCAATGAGAAAGCTCCAAATTTGATTTGGAATGAAA 720
DB 929 GAACAGCAAGATTAATTAATGATGATGCAATGAGAAAGCTCCAAATTTGATTTGGAATGAAA 988
QY 721 AATGGAATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGGATTTGTATGGGA 780
DB 989 AATGGAATTTCAAAAGCAAAAGAAAGAAAGCTGTGAAGACATTTGGGGATTTGTATGGGA 1048
QY 781 GTTTTCTATATATGCTGTGCTCTTTCTTATCTGTACAGTCATGACCTTTTCTTAC 840
DB 1049 GTTTTCTATATATGCTGTGCTCTTTCTTATCTGTACAGTCATGACCTTTTCTTAC 1108
QY 841 TACATTAATTCACCTTACTTTGAAATGATGT 869
DB 1109 TACATTAATTCACCTTACTTTGAAATGATGT 1137
```

Search completed: February 12, 2005, 00:52:33
Job time : 468 secs

Db 434 GTTATTGTGATGATCTTCAATAGTAGAGCTGCCGTCTTTTGGCAATTTGGAAATG 493
QY 481 ATCTTTCTGAGCTAACTTCAAGGCGCTGAAGAGATATATATACAAACATGTCACATGC 540
Db 494 ATCTTTCTGAGCTAACTTCAAGGCGCTGAAGAGATATATATACAAACATGTCACATGC 553
QY 541 AGAGAGAGTGTCTGTCTCTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCT 600
Db 554 AGAGAGAGTGTCTGTCTCTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCT 613
QY 601 TTTTATATACCTGATCTATTTATGTTATGTCTATATACAAATATATTTATCGCTTAA 660
Db 614 TTTTATATACCTGATCTATTTATGTTATGTCTATATACAAATATATTTATCGCTTAA 673
QY 661 GAACAGGCAAGATTAATAGATGATGCC-AATCAGAACTCCAAATTTGATGAAATGAA 719
Db 674 GAACAGGCAAGATTAATAGATGATGCCAAATGCAAGCTCCAAATTTGATGAAATG 733
QY 720 AAATGGAATTTCAAGCAAGCAAGAAAGAGAGCTGTGAAGACATTTGGGATTTGATGG 779
Db 734 AAAAATGGAATTTTCAACCAAGCAAAANAAAGGAAAGCTTGTGAAGACATTTGGG 793
QY 780 AGTTTCTCAATATAGTGTGTGCTCTTTCTTTATCTGTACAGTCATGACCTTTTCTT 837
Db 794 ATTGTGTGATGGGCAAGATTTTCCCTAAATATGCTGCGGTGGCCCTTTTCTT 851

RESULT 3

CN832845 824 bp mRNA linear EST 02-JUN-2004

LOCUS CN832845 15669803 NIH_MGC_145 Homo sapiens cDNA clone

DEFINITION IMAGE:30706626 5', mRNA sequence.

ACCESSION CN832845

VERSION CN832845.1 GI:47936624

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 824)

NIH-MGC http://mgs.mci.nih.gov/

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgsapbs-r@mail.nih.gov

Tissue Procurement: GPCR Consortium

cDNA Library Preparation: GPCR Consortium

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: IRB14 row: f column: 07

High quality sequence step: 519.

Location/Qualifiers

1..824

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30706626"

/tissue_type="mixed"

/lab_host="DH10B"

/clone_lib="NIH_MGC_145"

/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:

varies by clone; ORFs were PCR-amplified and cloned into

pCDNA3.1 by the GPCR Consortium. Cloning sites vary by

clone and include the following: 5'-EcoRV-Xmi1/Xhoi-3',

5'-EcoRV-Xmi1/NotI-3', EcoRV (TA cloned, non-directional).

For information about which gene each clone represents,

please visit our anonymous ftp site at

ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC Library."

Query Match 66.6%; Score 690.8; DB 7; Length 824;

Best Local Similarity 97.1%; Pred. No. 3.6e-165;

Matches 736; Conservative 0; Mismatches 17; Indels 5; Gaps 3;

ORIGIN

QY 1 TCAGGAATGATGACCTTTGGCACAATATTAATATATTTCTGTGTGAAAAACAATCG 60
Db 14 TCAGGAATGATGACCTTTGGCACAATATTAATATATTTCTGTGTGAAAAACAATCG 73
QY 61 TCAATATGATGTCGTCCTCTCCCTGACATTTAATGTCCTCATATTTCAACACATC 120
Db 74 TCAATATGATGTCGTCCTCTCCCTGACATTTAATGTCCTCATATTTCAACACATC 133
QY 121 GTTGGCAATCTGATATGTTATTTCTATATATCACTTCAACACATTTCAATCCCA 180
Db 134 GTTGGCAATCTGATATGTTATTTCTATATATCACTTCAACACATTTCAATCCCA 193
QY 181 AATTGCTCATTCATTCATGACGACATGACATTTCTTCTGGGGTGTCTGTATGCT 240
Db 194 AATTGCTCATTCATTCATGACGACATGACATTTCTTCTGGGGTGTCTGTATGCT 253
QY 241 TACAGTATGATGATGATCTGTGACGACATGTTGATATTTGAGAAAGCTCTGTAAAT 300
Db 254 TACAGTATGATGATGATCTGTGACGACATGTTGATATTTGAGAAAGCTCTGTAAAT 313
QY 301 CACACAGACCGACATTTATGTCGACGTCAGCTTCATTTTCCATTTTCTTCAATCTCC 360
Db 314 CACACAGACCGACATTTATGTCGACGTCAGCTTCATTTTCCATTTTCTTCAATCTCC 373
QY 361 ATTGACCGCTATGATGTCGTGTGATGATCACTGAGATTTAAAGCCAGATGATCTTG 420
Db 374 ATTGACCGCTATGATGTCGTGTGATGATCACTGAGATTTAAAGCCAGATGATCTTG 433
QY 421 GTTATTGTGATGATCTTCAATTAATGATGATGTCCTGCTGTTTGGCAATTTGGA 480
Db 434 GTTATTGTGATGATCTTCAATTAATGATGATGTCCTGCTGTTTGGCAATTTGGA 493
QY 481 ATCTTTCTGAGCTAACTTCAAGGCGCTGAAGAGATATATATACAAACATGTCACATGC 540
Db 494 ATCTTTCTGAGCTAACTTCAAGGCGCTGAAGAGATATATATACAAACATGTCACATGC 553
QY 541 AGAGAGAGTGTGCTGTCTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCT 599
Db 554 AGAGAGAGTGTGCTGTCTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCT 613
QY 600 TTTTATATACCTGATCTATTTATGTTATGTCTATATACAAATATATTTATCGCTTAA 659
Db 614 TTTTATATACCTGATCTATTTATGTTATGTCTATATACAAATATATTTATCGCTTAA 673
QY 660 AGAAGAGCAAGATTAATAGATGATGCCAAATGCAAGCTCCAAATTTGG-ATTGAAATGA 718
Db 674 AGAAGAGCAAGATTAATAGATGATGCCAAATGCAAGCTCCAAATTTGGAAATGA 733
QY 719 AAAATGGAATTT---CACAAAGCAAGAAAGAAAGCT 753
Db 734 AAAAAGGAATTTTCCCAAAACCAAAAGAAAGAAAGCT 771

RESULT 4

AY406936 675 bp DNA linear GSS 12-DEC-2003

LOCUS AY406936 675 bp DNA linear GSS 12-DEC-2003

DEFINITION Pan troglodytes HCM2723 gene, VIRUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

ACCESSION AY406936

VERSION AY406936.1 GI:39762907

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

	REFERENCE	1 (bases 1 to 675)	
	AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A., Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smunsky,J.J., Adams,M.D. and Cargill,M.	
	TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trios	
	JOURNAL	Science 302 (5652), 1960-1963 (2003)	
	PUBMED	14671302	
	REFERENCES	2 (bases 1 to 675) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A., Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smunsky,J.J., Adams,M.D. and Cargill,M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment.	
	COMMENT	Location/Qualifiers	
	FEATURES	1..675 <code>/organism="Pan troglodytes"</code> <code>/mol_type="genomic DNA"</code> <code>/db_xref="taxon:9598"</code> <code><1..2675</code> <code>/locus_tag="HNCM2723"</code>	
	ORIGIN		
	Query Match	63.3%; Score 657; DB 9; Length 675;	
	Best Local Similarity	97.8%; Pred. No.1.5e-156;	
	Matches	660; Conservative 0; Mismatches 15; Indels 0; Gaps 0	
OY	352	TTGATCTCCATTGACCGGTACTAGTGGTGTCGTGCATGCACTGAGATATAAAGCCAAGATG	411
Db	1	TTGATCTCCNNNAACCGGTACTANNNNNTGTGGATGCCACTGAGATATAAACCCAAGATC	60
OY	412	AATATCTGGTATTGTGGTAGATCTTCATTAGTTGAAGTGTCCCTGCTGTTTTGCA	471
Db	61	AATATCTTGTTATTTGTGTGATGATCTTCATTAGTTGAAGTGTCCCTGCTGTTTTGCA	120
OY	472	TTTGGATGATCTTTCTGAGCTAAACTCAAAGCGCTGAAGAATATATTACAACAT	531
Db	121	TTTGGATGATCTTTCTGAGCTAAACTCAAAGCGCTGAAGAATATATTACAACAT	180
OY	532	GTTCACTGAGAGGAGGTGCTGCTGCTCTTTCAGCAAAATATCTGGGGTAAGACTTT	591
Db	181	GTTCACTGAGAGGAGGTGCTGCTGCTCTTTCAGCAAAATATCTGGGGTAAGACTTT	240
OY	592	ATGACTTCTTTTATATACCTGATCTATATGATATGTCATATTAAGAAATATATCTT	651
Db	241	ATGACTTCTTTTATATACCTGATCTATATGATATGTCATATTAAGAAATATATCTT	300
OY	652	ATGCGTTAAAGAACAGGCAAGATTATATGTGATGCCAATCAGAAAGCTCCAATTGGATTG	711
Db	301	ATGCGTTAAAGAACAGGCAAGATTATATGTGATGCCAATCAGAAAGCTCCAATTGGATTG	360
OY	712	GAAATGAAAAATGGAATTTCACAAAAGCAAGAAAGAAAGCATGGAAGACATGGGATTT	771
Db	361	GAAATGAAAAATGGAATTTCACAAAAGCAAGAAAGAAAGCATGGAAGACATGGGATTT	420
OY	772	GTCATGGGAGTTTCTTAATATGCTGTGCTCTTTCTTATCTGTACAGTCATGACCCT	831
Db	421	GTCATGGGAGTTTCTTAATATGCTGTGCTCTTTCTTATCTGTACAGTCATGACCCT	480
OY	832	TTTCTTCGCTACATTATTCGACACTTGAATGATGTGGATTGTGGTTGGCTACTTG	891
Db	481	TTTCTTCGCTACATTATTCGACACTTGAATGATGTGGATTGTGGTTGGCTACTTG	540
OY	892	AACCTCACATTTAATCCAAATGTTTATGCAATTTTCTATCCTTGGTTAGAAAAGCACTG	951
Db	541	AACCTCACATTTAATCCAAATGTTTATGCAATTTTCTATCCTTGGTTAGAAAAGCACTG	600
OY	952	AAGATGATGCTGTTGGTAAAAATTTTCCAAAAGATTCATCCAGGTGTAATATTTTGG	1011

QY	Db	RESULT 5 AY406937 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL PUBLISHED REFERENCE AUTHORS TITLE JOURNAL COMMENT FEATURES ORIGIN	AY406937 Mus musculus HCM2723 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence. AY406937 GI:39762908 GSS. Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 927) Clark,A.G., Gnanowsk, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tandenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M. Interfering nonneutral evolution from human-chimp-mouse orthologous gene trios Science 302 (5652), 1960-1963 (2003) 2 (bases 1 to 927) Clark,A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tandenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment. Location/Qualifiers 1..927 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>927 /locus_tag="HCM2723"	601 AAGATGATGCTGTTTGGTAAATTTCCAAAAAGATTATTCAGGATGTAATTAATTTTGG 1012 GAATTGAGTTCAATG 1026 661 GAATTGAGTTCAATG 675	58.0%; Score 601.8; DB 9; Length 927; Best Local Similarity 78.7%; Pred. No. 1.9e-142; Matches 733; Conservative 0; Mismatches 192; Indels 6; Gaps 1;	94 ATGATGCTCATTAATTTGACACCACTGCTGGCAATCTGATAGTATTTGTTTATATCA 1 ATGTCACCTCAATTAATCCCGCCACACTCTGGTTGGCAATTAATAGTAATTAATTTCCATATCC 154 CACTTCAAAACAACCTTCATACCCCAACAATTTGGCTATTATTCCATGCGCACTGAGAC 61 CATTTCAAGCACTTCAATACCAACCAACACTGGCTCTTCACTCCAGGCGCATTTGTCAC 214 TTTCTTTCGAGGAGTCTGATGATGCTTACAGTATGATGATCTGCTGACACTGTGG 121 TTTCTGCTGGGAGTCTGATTAATGACCTGAGCATGATGTAAGAACTGTGAGCCGTGG 274 TATTTTGGAAAGTCTTCTGTAAATTTACACAAGACCGCAATTAATGCTGAGCTCAGCC 181 TATTTTGGGAAATTCCTGTGTAAGTTACACACGAGACCGATATCATGCTGAGCTCGCC 334 TCCATTTTCATTTGCTTTCATCTCCATGACCGGCTATCATGCTGTGTATTCACAGT 241 TCCATTTTCCACTTAGCTTTCACTTTTCATTTGACCGGCTAGCTGTGTGTGACCCCTTGG 394 AGATATTAAGCCCAAGATGAATATCTTGGTTATTTGTGTGATGATCTTCAATTAATGAGAT	453
----	----	---	---	---	---	---	-----

Db	301	AGATCAAAAGCCAAATCAATATCTCCACTATTTCTTGATATCTCTGTTAGTTGGAC	360
QY	454	GTCCCTGCTGTTTTCGATTTTGGAAATGATCTTTTCGAGACTAACTTCAAAGCGCTGAA	513
Db	361	CTTCTGCTGTTTATGCAATTTTGGATGATCTTCTTGAACTGAACCTTAAGAGAGTGGAA	420
QY	514	GAGATATATTACAAACATGTTCACTGCAGAGAGAGTGTCTGTCTTCTTTAGCAAAATA	573
Db	421	GAGCTGATTCGAGTCAGGTCAAGCAGCACTGGGCGGCTGTCTCCCTTTTATGTAAGTA	480
QY	574	TCGGGGTACTGACCTTTATGACTCTTTTATATACCGGATCATTAATGTTATGTATGTC	633
Db	481	TCGGGGTACTGACCTTTATGACTCTTTTATATACCGGATCATTAATGTTATGTATGTC	540
QY	634	TATTACAGAAATATATCTTATGCTTAAGAACAGGCAAGATTATATAGTATGCCAATCAG	693
Db	541	TACTATAGATATATATTTCACTAAGTAAAGGACAAAGCAAGGTCAATCATGTCAC-----G	594
QY	694	AAAGCTCCAAATTTGATTTGGAAATGAAAAATGGAATTTTCAACAAGCAAAAGAAAGAACT	753
Db	595	AATGTTCAAGTGTGATTTGAAAGGAAAAAGCCAAAGCACCAAAAGCAAGGAAAACAAAAGCG	654
QY	754	GTAAGAATATGGGGATTTGTATGGAGATTTTCCATATATGCTGTGTCCTTTCTTATAC	813
Db	655	GCGAAGACCTTAAAGGATCATGTGTGGCGCTTTCCTGTATGTGTGTGTCCTTTCTTC	714
QY	814	TGTACAGTCATGAGACCCCTTTTCTTCACTACATTTATTCACCTACTTTGATGATGTGTG	873
Db	715	TGCAGGCTCTGTGACCTTTTCTGTGGCTATGTTATCCACCTCTCTGATATACGCACTG	774
QY	874	ATTTCGTTGGCTACTGTAACCTACATCAATTTAATCCAAATGTTATGATGATTTTCTATCT	933
Db	775	TATTCGTTGGCTACTGTAATTTCTCCCTCAATCCGATGTGTTATGCTTTTCTATCCC	834
QY	934	TGCTTTAGAAAAGCACTGAAGTGAATGCTTGTGTTAATTTTCCAAAAAGATCATCC	993
Db	835	TGCTTTAGAAAAGCACTGAAGTGAATGCTTGTGTTAATTTTCCAAAAAGATCATCT	894
QY	994	AGGTGTAATTTATTTTTCGAAATTTGAGTTTCAT	1024
Db	895	AGGTGTAATTTATTTTTCGAAATTTGAGTTTCAT	925

RESULT 6

CO957870/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

789 bp

mRNA

linear

EST 17-AUG-2004

AGNCOURT 30842301 NIH_MGC_146 Homo sapiens cDNA clone

IMAGE:389847 3', mRNA sequence.

CO957870

CO957870.1 GI:51322452

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 789)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-remail.nih.gov

Tissue Procurement: Guthrie cDNA Resource Center

cDNA Library Preparation: Guthrie cDNA Resource Center

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/ILNL at:

<http://image.llnl.gov>

Plate: IRB18 row: 9 column: 05

FEATURES	
High quality sequence start: 24	
High quality sequence stop: 579.	
Location/Qualifiers	
source	1. .789

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7389847"
/tissue_type="mixed"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_146"
/notes="vector: pCDNA3.1; Site_1: multiple; Site_2:
multiple; ORFs were PCR-amplified (from IMAGE clones or
from commercially available cDNA libraries) and cloned by
the Guthrie cDNA Resource Center (www.guthrie.org/cDNA)
into pCDNA3.1. For specific information on cloning sites
(which vary by clone), please refer to the Guthrie
website, using the Guthrie ID given in the file
http://image.llnl.gov/image\_rearrayed\_plates/IRBP\_presv.data
a. Note: this is a NIH_MGC Library."
```

Query Match	53.3%	Score 553.2	DB 7	Length 789
Best Local Similarity	99.1%	Pred. NO. 4	7e-130	
Matches 577	Conservative 0	Mismatches 3	Indels 2	Gaps 2

OY		432	GATGATCTTCATTGGTGGAGTGCCCGCAGCTGTTTTGGCACTTGGAAATGACTTCTCTGA	491
Db		614	GATGATCTTCATTGG -TGGAGTGCCCTGCG -TTTGGCACTTGGAAATGACTTCTCTGA	557
OY		492	GCTAACTTCAAAGCGCCTGAGAAGATATATTAACAAATGTTCACTGCAGAGAGGTTG	551
Db		556	GCTAACTTCAAAGCGCCTGAGAAGATATATTAACAAATGTTCACTGCAGAGAGGTTG	497
OY		552	CTCGTCTTCCTTAGCAAAAATATCTGGGGATCTGACCCTTATGACTTCTTTTATATAC	611
Db		496	CTCGTCTTCCTTAGCAAAAATATCTGGGGATCTGACCCTTATGACTTCTTTTATATAC	437
OY		612	TGATCTATTATGTATATGTGTCTATTACAGAAATATATCTTATGCTAAGAACAGCAAG	671
Db		436	TGATCTATTATGTATATGTGTCTATTACAGAAATATATCTTATGCTAAGAACAGCAAG	377
OY		672	ATTATATTAGATGCCCAATCAGAGCTCCAAATTGGAAATGGAAAAATGGAATTC	731
Db		376	ATTATATTAGATGCCCAATCAGAGCTCCAAATTGGAAATGGAAAAAATGGAATTC	317
OY		732	ACAAGCAAAGAAAGAAAGCTGTGAAGACATTGGGGATTGTGANTGGAGTTTTCTAAT	791
Db		316	ACAAGCAAAGAAAGAAAGCTGTGAAGACATTGGGGATTGTGANTGGAGTTTTCTAAT	257
OY		792	ATGCTGATGCCCTTCTTTATCTGTACAGTCATGAGCCCTTTTCTTCACTACATTATTC	851
Db		256	ATGCTGATGCCCTTCTTTATCTGTACAGTCATGAGCCCTTTTCTTCACTACATTATTC	197
OY		852	ACCTACTTGAATGATGTGATTGGTTGGTGGCTACTTGAACCTCATCATTTAATCCAAT	911
Db		196	ACCTACTTGAATGATGTGATTGGTTGGTGGCTACTTGAACCTCATCATTTAATCCAAT	137
OY		912	GGTTATGACATTTTCTATCCTTGGTTAGAAAAAGCACTGAAGATGATGCTTTGGTAA	971
Db		136	GGTTATGACATTTTCTATCCTTGGTTAGAAAAAGCACTGAAGATGATGCTTTGGTAA	77
OY		972	AATTTGCCAAAAAGATCATCCAGGCTGAATAATTATTTTGGAA 1013	
Db		76	AATTTGCCAAAAAGATCATCCAGGCTGAATAATTATTTGGAA 35	
<hr/>				
RESULT 7				
CO957854				
LOCUS				
DEFINITION				
IMAGE:				
ACCESSION				
<hr/>				
CO957854				
ABENICOURT 30842669 NIH_MGC_146 Homo sapiens cDNA clone				
IMAGE:7389847 5' , mRNA sequence.				
CO957854				

VERSION C0957854.1 GI:51322436
 EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 684)
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Guthrie cDNA Resource Center
 cDNA Library Preparation: Guthrie cDNA Resource Center
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

FEATURES
 source
 High quality sequence stop: 460.
 Location/Qualifiers
 1..684

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7389847"
/issue_type="mixed"
/lab_host="DH10B (T1-phage-resistant)"
/notes="Vector: pCDNA3.1; Site 1: multiple; Site 2: multiple; ORF's were PCR-amplified (from IMAGE clones or from commercially available cDNA libraries) and cloned by the Guthrie cDNA Resource Center (www.guthrie.org/cDNA) into the pCDNA3.1. For specific information on cloning sites (which vary by clone), please refer to the Guthrie website, using the Guthrie ID given in the file ftp://image.llnl.gov/image.rearrayed_plates/IRBF.presv.dat
a. Note: this is a NIH_MGC library."

```

Query Match 49.0%; Score 508.4; DB 7; Length 684;
 Best Local Similarity 96.9%; Pred. No. 1.2e-118;
 Matches 526; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

```

QY 7 ATGATGCCCTTTGGCACAATATTAATATTTCTGTGGAGAAAACAAGCGGCAAT 66
DB 2 ATGATGCCCTTTGGCACAATATTAATATTTCTGTGGAGAAAACAAGCGGCAAT 61
QY 67 GATGCCGTCCTTCCCTGTACAGTTAAAGTGTCATTAATTCGACACACCTCGTGGC 126
DB 62 GATGCCGTCCTTCCCTGTACAGTTAAAGTGTCATTAATTCGACACACCTCGTGGC 121
QY 127 AATCGATAGTATTTGTTCTATATCACACTTCAACCACTTACCCCAAAATTTGG 186
DB 122 AATCGATAGTATTTGTTCTATATCACACTTCAACCACTTACCCCAAAATTTGG 181
QY 187 CTCATTCATTCAGGACCTGTGAGACTTTCTTCGGGGTGTCTGTCATCTTACAGT 246
DB 182 CTCATTCATTCAGGACCTGTGAGACTTTCTTCGGGGTGTCTGTCATCTTACAGT 241
QY 247 ATGGGAATATGTCGAGACCTGTTGGATTTTGAAGAAGCTTCTGTAATTCACACA 306
DB 242 ATGGGAATATGTCGAGACCTGTTGGATTTTGAAGAAGCTTCTGTAATTCACACA 301
QY 307 AGCACCAGCATTAATGAGCTCAGCTCACTTTTCCATTTGTTCACTTCATTCAGTAC 366
DB 302 AGCACCAGCATTAATGAGCTCAGCTCACTTTTCCATTTGTTTTCATCTCCATTGAC 361
QY 367 CGCTACTATGCTGTGTGATCCACTGAGATATAAAGCAAGATGATATCTTGTTATT 426

```

```

DB 362 CCGTACTATGCTGTGTGATCCACTGAGATATAAGCAAGATGATATCTTGTTATT 421
QY 427 TGTGATGATCTTCATTAGTTGAGTGTCCTCGTGTGTTTGCATTTGGAATGATCTTT 486
DB 422 TGTGATGATCTTCATTAGTTGAGTGTCCTCGTGTGTTTGCATTTGGAATGATCTTT 481
QY 487 CTGAGACTTAACTTCAAGGCGCTGAGAGATATATTCAAAACATGTTCACTGCAGAGA 546
DB 482 CTGAGACTTAACTTCAAGGCGCTGAGAGATATATTCAAAAC-TGTTGCTTTAGAGA 540
QY 547 GGT 549
DB 541 GGT 543

```

RESULT 8
 CN838940/c
 LOCUS CN838940.1
 DEFINITION AGNCOURT.15864245 NIH_MGC.145 Homo sapiens cDNA clone
 IMAGE:30706625 3', mRNA sequence.
 ACCESSION CN838940
 VERSION CN838940.1 GI:47944595
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 929)
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-remail.nih.gov
 Tissue Procurement: GPCR Consortium
 cDNA Library Preparation: GPCR Consortium
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

Query Match 41.6%; Score 431.4; DB 7; Length 929;
 Best Local Similarity 98.5%; Pred. No. 5.9e-99;
 Matches 446; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

FEATURES
 source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706625"
/issue_type="mixed"
/lab_host="DH10B"
/clone_lib="NIH_MGC_145"
/notes="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORF's were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmiI/XhoI-3', 5'-EcoRV-XmiI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/rearrayed_plates/IRBF.presv.dat
a Note: this is a NIH_MGC library."

```

Query Match 41.6%; Score 431.4; DB 7; Length 929;
 Best Local Similarity 98.5%; Pred. No. 5.9e-99;
 Matches 446; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```

QY 587 CCTTATGACTTC-TTTTATATACCTGAGATCTATTATGATGCTGCTTATTAAGATA 645
DB 752 CCTTATGACTTCCTTTTATATACCTGAGATCTATTATGATGCTGCTTATTAAGATA 693

```


QY 646 TATCTATCGCTAAAGAACGAGATTAAATGATGATCCCAATCAGAGCTCCAAATT 705
| | | | |
692 TATCTATCGCTAAAGAACGAGATTAAATGATGATCCCAATCAGAGCTCCAAATT 633
| | | | |
QY 706 GGATTGGAATGAAATGAAATTTCCAAAGCAAGAAAGAAAGCTGTGAACATTTG 765
| | | | |
Db 632 GGATTGGAATGAAATGAAATTTCCAAAGCAAGAAAGAAAGCTGTGAACATTTG 573
| | | | |
QY 766 GGGATTGATGGAGTTTCTTAATATGCTGTCCTTTTATCTGTACATGATG 825
| | | | |
Db 572 GGGATTGATGGAGTTTCTTAATATGCTGTCCTTTTATCTGTACATGATG 513
| | | | |
QY 826 GACCTTTTCTTCACTACATTAATTCACCTACTTGAATGATGTGTTGGTTGGC 885
| | | | |
Db 512 GACCTTTTCTTCACTACATTAATTCACCTACTTGAATGATGTGTTGGTTGGC 453
| | | | |
QY 886 TACTGAACTGACATTTAATCAATGTTTATGCAATTTTCTGTGTTAGAAA 945
| | | | |
Db 452 TACTGAACTGACATTTAATCAATGTTTATGCAATTTTCTGTGTTAGAAA 393
| | | | |
QY 946 GCACGAGATGATGCTGTTTGTGTAATAATTTCCAAAAAGATTCATCCAGGTGAATTA 1005
| | | | |
Db 392 GCACGAGATGATGCTGTTTGTGTAATAATTTCCAAAAAGATTCATCCAGGTGAATTA 333
| | | | |
QY 1006 TTTTGAATTGATTCATAGATTATTTATTT 1038
| | | | |
Db 332 TTTTGAATTGATTCATAGATTATTTATTT 300
| | | | |

RESULT 9
CN832787/c 941 bp mRNA linear EST 02-JUN-2004
LOCUS CN832787.1 15864261 NIH_MGC_145 Homo sapiens cDNA clone
DEFINITION IMAGE:30706626 3', mRNA sequence.
ACCESSION CN832787
VERSION CN832787.1 GI:47936566
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 941)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: GPCR Consortium
cDNA Library Preparation: GPCR Consortium
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Place: IRB14 row: E column: 07
High quality sequence scop: 646.
Location/Qualifiers
1..941
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706626"
/tissue_type="mixed"
/lab_host="DH10B"
/clone_lib="NIH_MGC_145"
/note="Vector: pCDNA3.1, Site 1: varies by clone; Site 2:
varies by clone; ORFs were PCR-amplified and cloned into
pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
clone and include the following: 5'-EcoRV-XbaI/XhoI-3',

5'-EcoRV-XbaI/XhoI-3', EcoRV (TA cloned, non-directional).
For information about which gene each clones represents,
please visit our anonymous ftp site at
http://image.llnl.gov/image/retarrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC library."

ORIGIN
Query Match 40.2%; Score 417; DB 7; Length 941;
Best Local Similarity 94.3%; Pred. No. 2,8e-95;
Matches 465; Conservative 0; Mismatches 25; Indels 3; Gaps 3;
QY 549 TTGCTGTGCTCTTTAGCAAAATATCTGGGACTGACCTTTATGAC-TTCTTTTATA 607
| | | | |
Db 795 TTTTCTCTTATAGCAAAATATATGAGGACTTACCTTTATGACTTTCTTTTCA 736
| | | | |
QY 608 TACCTGAT-CTATTATGTATGTCTCTTATACAAATATATCTTATGCTTAAACAG 666
| | | | |
Db 735 TCCCTGATGCTCTTATGTATGTATTTATTTATTAACAATATTTCTTATGTTAAACAG 676
| | | | |
QY 667 GCAAGATTAAATTAGATGATCCATCAGAAAGCTCCAAATTGATGGAATGAAATGGA 726
| | | | |
Db 675 GCAAGATTAAATTAGATGATCCATCAGAAAGCTCCAAATTGATGGAATGAAATGGA 616
| | | | |
QY 727 ATTTCACAAAGCAAGAAAGAAAGC-TGTGAAGACATTGGGATGTGATGGAGTTT 785
| | | | |
Db 615 ATTTCACAAAGCAAGAAAGAAAGC-TGTGAAGACATTGGGATGTGATGGAGTTT 556
| | | | |
QY 786 CCTATATGCTGTGTCCTTTCTTTATCTGTACATGACCTTTTCTTCTACTACT 845
| | | | |
Db 555 CCTATATGCTGTGTCCTTTCTTTATCTGTACATGACCTTTTCTTCTACTACT 496
| | | | |
QY 846 TATTCACCTACTTTGATGATGATGTTGATTTGTTGGCTACTTGAATCTACATTTTA 905
| | | | |
Db 435 TATTCACCTACTTTGATGATGATGTTGATTTGTTGGCTACTTGAATCTACATTTTA 436
| | | | |
QY 906 TCCAAATGTTATGATTTTCTATCTGTTTGAAGAAAGCACTGAAGATGATCTGTT 965
| | | | |
Db 435 TCCAAATGTTATGATTTTCTATCTGTTTGAAGAAAGCACTGAAGATGATCTGTT 376
| | | | |
QY 966 TGTGTAATTTTCCAAAGAAATTCATCCAGGTGAATTTTGGATGATGATCTTA 1025
| | | | |
Db 375 TGTGTAATTTTCCAAAGAAATTCATCCAGGTGAATTTTGGATGATGATCTTA 316
| | | | |
QY 1026 GAATTATTTATTT 1038
| | | | |
Db 315 GAATTATTTATTT 303
| | | | |
RESULT 10
CNS0407J/c 1016 bp DNA linear GSS 01-SEP-2000
LOCUS CNS0407J.1 129C03 of library G from Tetradon nigroviridis, genomic survey
DEFINITION sequence.
ACCESSION AL302248
VERSION AL302248.1 GI:8180344
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neocleosteii;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1
Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fzames,C., Wincker,P., Brothier,P., Quetler,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetradon nigroviridis DNA sequence
Nat.Genet. 25 (2), 235-238 (2000)
JOURNAL MEDLINE
20296633
REFERENCE 10835645

AUTHORS
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL
Genome Res. 10 (7), 939-949 (2000)
MEDLINE
20359837
10899143
PUBMED
3 (bases 1 to 1016)
REFERENCE
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
JOURNAL
- Web : www.genoscope.cns.fr
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
1..1016
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone_1b="Tg"
/note="Genoscope sequence ID : C0BG129AB02SP1-end :
PUC-Or1"
ORIGIN
Query Match 25.3%; Score 262.8; DB 9; Length 1016;
Best Local Similarity 55.2%; Pred. No. 6.4e-56;
Matches 544; Conservative 0; Mismatches 423; Indels 18; Gaps 1;
QY 34 AATATTCCGTGTGTAACAACTGTCATAATGATGTCGGTCTCCCTGACAGTTTA 93
DB 967 AACATTCTCTGTGTAAAGACCGCTATTCACAGCCCTTCGTGTATCTTTACATGGTT 908
QY 94 ATGGGCTCATTAATCTGACCACTCGTGGCAATGTGATGTTATGTTCTATATCA 153
DB 907 TTAGGGGTATGTGATGTCGTCGTGTGGAACCTTTGGTCACTTTCTATCANC 848
QY 154 CACTTCAACAACCTTCATCCCAACAAATGGCTCATTCATTCATGAGCCACTGTGAC 213
DB 847 TATTTTAAAGCACTCACTCAACCACTCACTGTTGCTCTCTGCTGTGTCTAC 788
QY 214 TTTCTTGGGGGTGTGTCATGCTTACAGTATGTGATGTGTCGAGCACTGTGG 273
DB 787 TTTCTCTGGGGGTGTGTCATGTTTACCAATATGATTCAGACATGAAACCTGCTGG 728
QY 274 TATTTTGAAGAAGTTCCTGTAATTCACACAGACCGACATTATGCTGAGCTGACC 333
DB 727 TATTTTGAAGACATTTTGGCAAGTTCATTGAGTTCTGATGTGATGTTGTGACAGCA 668
QY 334 TCCATTTTCAATTTGCTTTCATCTCCATTGACCGCTACTATGCTGTGTGATGCCACTG 393
DB 667 TCCATTAATTAATGCTCTTAAATATGATGATGATGATGATGATGATGATGATGATGATG 608
QY 394 AGATTAAGCAAGATTAATATCTGTTATTTGTGTGATGATCTTCATTAGTTGAGT 453
DB 607 GTGTATTAAGAAAAATGACTGTGATGATGTTGTTGTGATGATGATGATGATGATGATG 548
QY 454 GTCCCTGTGTTTGTGATTTGGAATGATCTTTGAGAGCTTAACCTCAAGAGCGCTGAA 513
DB 547 GTTTCAGGTGTGTGATGTTTGAATGATTTTAAAGCTGAACATTTGGGAATTTGAA 488
QY 514 GAGATTAATTAACAACATGTTCACTGCAAGAGAGTGTCTGTCTCTTTTCAAAATA 573
DB 487 GATTTTCAACAACATTTGTGATGGAAGAGATGATTTTGTGACAGAGGTCG 428
QY 574 TCTGGGGTACTGACCTTATGACTCTTTTATATACCTGATCTATTAATGTTATGTGTC 633
DB 427 TCAAGTACAGTCTATGATGCTCTGCTTTTATATTCAGGAATTAATCATGATCAGTTTA 368

QY 634 TATTACAGAAATATATTATTCCTTAAGAACAGCAAGATTATATTGATGCCAATCAG 693
DB 367 TAGCGAAGATTTCTTTGTGTGCAAGAGACAAATTCCTCAGATTCAGAAATCAAGCAGC 308
QY 694 AGCTCCAAATTTGATTGGAATGAAAATGGAATTTCACAAAGCAAGAAAGAAAGCT 753
DB 307 ATG-----ACATCAGCCAGAGATTCAGTAAAGAAAAACAGCAAAAGCC 266
QY 754 GTGAAGACATTTGGGATTTGTATGAGAGATTTTCCATATATGCTGTGCTTTCTTATC 813
DB 265 ACTAAACCTGTTGTGTATCATGAGACATTTCTTTATGCTGAGCTCAGTTTCTC 206
QY 814 TGTAAAGTATGAGACCTTTTCTTCACTACATTAATTCACCTACTTTGAATGATGTG 873
DB 205 TGTTCATCATTTGATCCCTTTTTCAGTACCAACCAACAGATGATGTTTGAACAATTC 146
QY 874 ATTTGGTGGCTCTTGAACCTTACATTAATTAATGATGTTATGATGATTTTCTATCCT 933
DB 145 CTATGGCTAGTTATATGAAATTTCACTAAACCTCTCATTTATGCAATTTTTCACACC 86
QY 934 TGGTTTGAAGAAAGCACTGAAGATGATGCTGTTTGTGTAATTTTCCAAAAAGATTCAATCC 993
DB 85 TGGTTTGAAGAGCGTTTTCCTGTTTGTCTTCAGAGAAAGCTTTAAAGAGATATGTCA 26
QY 994 AGCTGTAAATTTATTTTGAATTTCA 1018
DB 25 AACATTTCCCTTTTACTTCATTA 1
RESULT 11
CNS05JP2 1015 bp DNA linear GSS 01-SEP-2000
LOCUS
Tetraodon nigroviridis genome survey sequence T7 end of clone
008G08 of library A from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION
AL340463.1 GI:8234221
VERSION
GSS: genome survey sequence.
KEYWORDS
Tetraodon nigroviridis
SOURCE
Tetraodon nigroviridis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE
1 Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fischer,C.,
Bernot,A., Fizames,C., Wincker,P., Brotilier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
JOURNAL
MEDLINE
20296633
PUBMED
10835645
REFERENCE
2
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
JOURNAL
MEDLINE
20359837
PUBMED
10899143
REFERENCE
3 (bases 1 to 1015)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers

```

source
1. 1015
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone_lib="A"
/note="Genoscope sequence ID : C0AA008B04C1-end : T7"

ORIGIN

Query Match      24.8%; Score 257.6; DB 9; Length 1015;
Best Local Similarity 55.9%; Pred. No. 1,4e-54;
Matches 510; Conservative 0; Mismatches 385; Indels 18; Gaps 1;

QY 34 AATATTTCTGCTGTAAGAAAACACATGTCATATGATGTCCTTCCCTGACAGTTA 93
DB 98 AACATTTCTGCTGTAAGAAAACCTCTATACATGCCCATTCGTGTCATCTTATACATGTT 157
QY 94 ATGTCCTCATATATCTGACCACTCGTTGGCAATCTGATAGTATGTTTCTATATCA 153
DB 158 CTAGGGGNCATGTTGTTGTCATGTCGTGTAAGAACTTTTGGTCACGTGTTCTATATC 217
QY 154 CACTTCAACCACTTCATACCCCAACAAATTGGCTCATTCATCCATGGCCACTGTGAC 213
DB 218 TATTTTAAAGCACTCCACCTCCAACTACCTGTTGCTCTCTGCTGTCGTGAC 277
QY 214 TTTCTTCTGGGTCGTCTGTCATGCTTACAGTATGTCATGTCATGTCATGTCATGTC 273
DB 278 TTTCTCTGGGTCGTCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 337
QY 274 TATTTTGAAGAGTCTTCTGTAATAATTCACACAGCACCAGATTAATGCTGACGCC 333
DB 338 TATTTTGAAGACATTTTGGCAAGATACATNNNGATTTGCTGNAATGTTGTGACAGCA 397
QY 334 TCCATTTTCCATTTGCTTTCATCTCCATGACCGCTACTATGCTGTGTGATCCACTG 393
DB 398 TCCATTAATTAATCTCTTAAATATCAATGATCAATAAGCGGTGATATCTCTCA 457
QY 394 AGATATTAAGCCAAAGATGATCTTGGTTATTTGTCATGATCTTCTGTTAGTGAAGT 453
DB 458 GGTGTTAAGAAAATGTCGTGTCATGTTGTTGTCATGATTTTGTGACGTGGGT 517
QY 454 GTCCCTGCTGTTTTCATTTGGAATGATCTTCTGAGACTTAACCTTCAAGCGCTGAA 513
DB 518 GTTTCAGGTCGTGTCATGTTGGAATGATTTTTCAGGCTGAATATTTTGGGAATGAA 577
QY 514 GAGATATATTAACAACATGTCATGTCAGAGAGAGTGTCTGTCTCTTTCGCAAAATA 573
DB 578 GATTTCTACAAATGACTTGTGTCATGTCAGAGAAATGATATTTGTTCAAGTGTATG 637
QY 574 TCTGGGTAAGTACCTTTATGACTTCTTTTATATACCTGATCTATTAATGTTATGTC 633
DB 638 TCAAGCAGAGTCTGTCAGTCTCTCTTTTATATTCGGAATATATGATGTTGTGTA 697
QY 634 TATTAACAGAAATTAATCTTATGCTTAAGAAACAGCAAGATTAATTAATGATGCCAATCAG 693
DB 698 TACCTGAAGATTTTCTTGTGCAAGAGAACATTCCTCAGCATTCAGAAATCAACCTGC 757
QY 694 AAGCTCCAAATGATGGAATGAATAATGAAATTTTCACAAAGCAAGAAAGAAAGT 753
DB 758 ATG-----ACATCAGCAGCAATTTCAAAATTAAGCAAGCAAGCAAGCAAGC 799
QY 754 GTGAGACATTTGGGATTTGTGATGGAGATTTTCTAATATGCTGGGCTCTTCTTTATC 813
DB 800 ACTAATAACGTGTTGTGTATCATGAGAGGCTTTCTTTATGCTGGGCTCAATTTTCTC 859
QY 814 TGTACAGTCATGAGCCCTTTCTTCACTACATTAATTCACCTTAATGATGATGTTG 873
DB 860 TGTATATGTCATGATCCTTTCACTACCTCTCAACACAGCTGATGTAAGAAATTT 919
QY 874 ATTGTCCTGCTGCTGATGAACTCTACATTAATCCAAATGTTTATGATTTTCTATCT 933
DB 930 GGATGGGNGGTTTATTAATTTCTACACAAACCCATCATTAATGATTTTATACAC 979

```

```

QY 934 TGGTTTACAAAAG 946
DB 980 TGGTTTACAAAAG 992

RESULT 12
AW813121/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 536)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.G.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?el=ct2-CM4-ST0189-051
099-021-h04&3=1989-10-05&4=1)
Seq primer: puc 18 forward
High quality sequence start: 26
High quality sequence stop: 534.
Location/Qualifiers
1. 536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ST0189"
/note="Organ: stomach; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match      23.9%; Score 248; DB 2; Length 536;
Best Local Similarity 98.0%; Pred. No. 3,4e-52;
Matches 251; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 783 TTTCTATATGCTGTCGTCCTTTCTTATATCTGACATGACCCCTTTCTTCACTA 842
DB 535 TTTCTATATGCTGTCGTCCTTTCTTATATCTGACATGACCCCTTTCTTCACTA 476
QY 843 CATTAATTCACCTACTTGAATGATGTTGATTTGTTGGCTACTGAACCTACACTT 902
DB 475 CATTAATTCACCTACTTGAATGATGTTGATTTGTTGGCTACTGAACCTACACTT 416
QY 903 TAATCAATGGTTTATGATTTTCTATCTTGGTTTGAAGAAAGCACTGAAGATGATCT 962

```

Db 415 TAAATCCATGTTTAAAGCATTTTCTATCCCTGGTTTGAAGAAAGCACTGAAGATGATGCT 356
QY 963 GTTGTGTAATATTTTCCAAAAGATTCATCCAGGTGAATATTTTGGATTTGATGCTC 1022
Db 355 GTTTGGTAAATTTTTCCTCAAAAGATTCATCCAGGTGAATATTTTGGATTTGATGCTC 296
QY 1023 ATAGAAATTTATTTATTT 1038
Db 295 ATAGAAATTTATTTATTT 280

RESULT 13
CD559484/c 797 bp mRNA linear EST 19-NOV-2003

LOCUS AGENCOURT 14496575 NIH_MGC_195 Homo sapiens cDNA clone
ACCESSION IMAGE:6971835 5', mRNA sequence.
CD559484
VERSION CD559484.2 GI:38453527

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS 1 (bases 1 to 797)
TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
On Jun 10, 2003 this sequence version replaced gi:31585552.

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@b3-remail.nih.gov

Tissue Procurement: Narayan Bhat
CDNA Library Preparation: Bhat Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMN)

DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILMN at:

http://image.llnl.gov
Plate: IRB82 row: e column: 02
High quality sequence start: 10
High quality sequence stop: 798.

FEATURES
location/Qualifiers
1..797
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6971835"
/tissue_type="mixed"
/lab_host="DH5A (TI phage-resistant)"
/clone_lib="NIH_MGC_195"
/note="Vector: pDNR-Dual; Site 1: loxp-sali; Site 2:
loxp-hindiii; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxp sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB8.presv.dat
a Note: this is a NIH_MGC Library."

ORIGIN

Query Match 20.5%; Score 213; DB 6; Length 797;
Best Local Similarity 57.6%; Pred. No. 3.1e-43;

Matches 437; Conservative 0; Mismatches 295; Indels 27; Gaps 2;
QY 220 CTGGGGTGTCTGTCATGCTTACAGTATGATGAGATGCTGAGACACTGTGGATTTT 279
Db 797 CTGATTTCCACATCATCATCATATGATGATGATGATGATGATGATGATGATGATGAT 738
QY 280 GGAAGAAGTCTCTGTAATTTCAACAAGCAACCAATTATGCTGAGCTCAGCTTCATT 339
Db 737 GGGCTTACATTTTGGAAATTTATTTATTTTACCGATGATGATGATGATGATGATGATGAT 678
QY 340 TTCCATTTGCTTTCATGCTTACATGACCGCTACTAGTGTGCTGCTGCTGCTGCTGCTGCT 399
Db 677 TTTCATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
QY 400 AAAGCCAAATGAAATATCTGTTATTTGCTGATGATGATGATGATGATGATGATGATGAT 459
Db 617 TCACCAAAATATATATATCCAGTCAATTAAGAATGCTTACCTTATGTTGGTGGCTCCT 558
QY 460 GCTGTTTTGATTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 519
Db 557 GAGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
QY 520 TATTACAAATGCTTTCATGCTGAGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579
Db 500 TATGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
QY 580 GTACTGACCTTTATGATCTTCTTTTATATACCTGATCTATATATATATATATATATATAT 639
Db 440 ACCACCTTGTATATGAGAGGCTTCTTCACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 381
QY 640 AGATATATCTTATTCGTTAAAGAAACAGGCAAGTATATATGATGATGATGATGATGATGAT 639
Db 380 AAAATTTTTCAGATATCAAAAACATGCTCAGTCCATCAATTA----- 337
QY 700 CAATTTGATTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 759
Db 336 -----CTTGGAGAGAAATCAAAATTAATCAAGTGAAGAAAGCAAAAAGCTGCCAAA 285
QY 760 ACAATGGGGAATTTGATGAGAGATTTTCTTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 819
Db 284 ACTTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 225
QY 820 GTATGAGACCTTTTCTTCTGCTACTATATTTCCACCTACTTGAATGATGATGATGATGATGAT 879
Db 224 TTATGATGATCCCTTTTGAATCTTCTACTCTGATGATTTTGTATATGCTTGAATGATGAT 165
QY 880 TTGGCTACTTGAATCTTCAATTTAATCCAAATGATTTATGATGATTTTCTATCTTGTGTT 939
Db 164 TTGGCTATTTTAACTCCACATGATATCCGTTAATATATGATGATTTCTTATCTTGTGTT 105
QY 940 AGAAAGCACTGAAGATGATGCTGTTGGTAAATTTTC 978
Db 104 CGCAGAGCACTGAAGATGATGATTTTGTCTAGTAAATTTTC 66

RESULT 14
CD559544 773 bp mRNA linear EST 26-NOV-2003

LOCUS AGENCOURT 14496987 NIH_MGC_195 Homo sapiens cDNA clone
DEFINITION IMAGE:6971759 5', mRNA sequence.
CD559544
ACCESSION CD559544.2 GI:3858970

VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 773)
TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
On Jun 10, 2003 this sequence version replaced gi:31585612.
Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabs-remail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: Bhat Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: IRBK1 row: f column: 10
High quality sequence stop: 741.
Location/Qualifiers

FEATURES
source

```
1..773
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6971759"
/tissue_type="mixed"
/lab_host="DH5A (T1 phage-resistant)"
/clone_1lb="NIH_MGC_195"
/notes="Vector: pDNR-Dual; Site_1: loxp-Sall; Site_2:
loxP-HindIII; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxp sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearranged_plates/IRBK.presv.dat
A Note: this is a NIH_MGC library."
```

ORIGIN

```
Query Match      20.1%; Score 208.4; DB 6; Length 773;
Best Local Similarity 56.8%; Pred. No. 4,6e-42;
Matches 423; Conservative 0; Mismatches 317; Indels 5; Gaps 2;

QY 14 CTTTTCGCACATATTAATTAATTTCTGTGTGAACAACTGCTGAATGATGCTC 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 34 CCAAGTTGTCAAATTTGTAAATTAATCTGTCTCCACCAACCGCTTTTCATGTC 93

QY 74 GTGCTTCCCTGTACAGTTAATGATGCTCATATTTCTGACACACTGTTGCAATCTGA 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 94 AGGTGATATGTTATGCGTTATGACGAGCCATGTTA--TCATATTGCGAAACTTGG 151

QY 134 TAGTATATGTTTCTATATACACTTGAACAACTTCATACCCCAACAAATGGCTATTC 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 152 TTAATATGTTTCCATATCGCATTTCAACAGCTTCACCTCCCAACAACTTTCATGATCC 211

QY 194 ATTCCATGGCCACTGTGACTTTCTCTGGGGTGTGTGTCATGCTTACATGATGATGA 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 212 TCTCATATGCAACCAACGACCTTCTGCTGGGTTTGTCTATTAATGACATACACATATATGC 271

QY 254 GATCTGCTGAGCACTGTGATATTTTGGAGAAGCTTCTGTAAATTTACACAAAGACCG 313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 272 GATCAGTGAAGAGTTCTGCTGTACTTTGGGGAATGGCTTTGTAAATTTCCACAAAGCTTTGG 331

QY 314 ACATTAATGCTGAGCTCAGCCTTCATTTTCATTTGCTTTCATCTTCATGACCGCTACT 373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 332 ACATGATGCTCAGACTGACCTTCATTTCCACTCTGTTCCATTTGCTATTAAGACGATTTT 391

QY 374 ATGCTGTGTGATCCACTGAGATATTAAGCCAAAGTGAATATCTTGGTTATTTGTGTA 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 392 ATGCGCTGTGTAACTTTACATTAACACCAACAAATGAAGAACTCCACCATTAAGAACAC 451
```

```
QY 434 TGATCTCATTAAGTTGAGATGTCCTGCTGTTTTTTCATTTGGAAATGATCTTTCTGAGC 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 452 TGCTGGCAATTTTGTGTCAGATTTCTGCTCTTTTCTTTTGGTTAGTTATCTGAGG 511

QY 494 TAACTTCAAGGCGCTGAAGATATTTACAAACATGTTTCACTGCGAGAGAGTTGCT 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 512 CCGAAGT---TTCCCGATGCGAGACTTAAATATCTGTTGCTTTCATATTTCTGTG 568

QY 554 CTGCTTCTTTAGCAAAATATCTGGGATACGACCTTTATGACTTCTTTTATATACCTG 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 CCTTACTTTTCAACAAATTTCTGGGGACATATTTGTTCACTACATGTTTCTTTACCCCTG 628

QY 614 GATCTATTATGTTATGTCATTAATTAACAGATATATCTTATCGCTTAAGAACGAGAT 673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 629 GCTTCATGATGTTGTGATTTATATGCAAAATCTTATGTTTCCAAACGATGCTCGAG 688

QY 674 TAATTAATGATGCCATATCAGAACTCCAAATTTGATTTGGAATGAAATGGAATTTTC 733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 689 TCATCAGCCATGTGCTTAAACACAAAGGGGGCAGTAAACACCTATCCMAAGAAAG 748

QY 734 AAAGCAAGAAAGAAAGAGCTGTGA 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 749 GACAGAAAGCAGCGACCTTGATA 773
```

RESULT 15
CNS04VLK
LOCUS
DEFINITION
CNS04VLK 1071 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence T3 end of clone
035005 of library A from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION
AL030923.1 GI:8216172
VERSION
GSS: genome survey sequence.
KEYWORDS
Tetraodon nigroviridis
SOURCE
ORGANISM
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorphi; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE
AUTHORS
Roest Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Queller,F.,
Saurin,W. and Weissenbach,J.
TITLE
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL
Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE
20296633
PUBMED
10835645

REFERENCE
AUTHORS
Roest Crollius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Queller,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL
Genome Res. 10 (7), 939-949 (2000)
MEDLINE
20359837
PUBMED
10899143

REFERENCE
AUTHORS
Genoscope.
TITLE
Direct Submision
JOURNAL
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
Web : www.genoscope.cns.fr

COMMENT
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
source

```
1..1071
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 02:33:13 / Search time 74 Seconds
(without alignments)
1771.781 Million cell updates/sec

Title: US-09-980-145-6
Perfect score: 1806
Sequence: 1 MNPFCNIIINISCVNNSN.....FGKIFKDSRCKLFLKLS 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1806	100.0	339	4	AAB49232 Human SNO
2	1806	100.0	339	4	AAG80973 Human ncp
3	1806	100.0	339	5	ABJ04073 Human G p
4	1806	100.0	339	5	ABG93791 Human G p
5	1806	100.0	339	5	ADH48916 NOV84B pr
6	1806	100.0	339	5	ADH48908 NOV81B pr
7	1806	100.0	339	5	ADH48906 NOV81A pr
8	1806	100.0	339	6	ABP81732 Human tra
9	1806	100.0	339	8	ADG09662 Human GPC
10	1806	100.0	339	8	ADG33691 Human tra
11	1801	99.7	338	4	ADG33691 Human tra
12	1791	99.2	339	5	ABH80694 Human tra
13	1539	85.2	311	7	ADG12792 Human GPC
14	1525	84.4	296	5	AAG80972 Human ncp
15	1525	84.4	296	5	ABG93790 Human G p
16	1403	77.7	332	4	AAB49231 Rat SNORF
17	1333	77.1	332	5	ABH80693 Rat trace
18	1351	74.8	332	4	AAB49234 Mouse SNO
19	1351	74.8	332	8	ADG09663 Mouse GPC
20	1212	67.1	238	4	AAG80960 Human ncp
21	1212	67.1	238	5	ABG93778 Human G p
22	998	55.3	191	4	AAB49230 Human SNO
23	923.5	51.1	347	8	ADG29145 Mouse nov
24	886.5	49.1	343	8	ADG29145 Mouse GPC
25	848	47.0	351	5	ABH79888 Human SNO

26	848	47.0	351	7	ADE81115
27	840.5	46.5	306	8	ADO29074
28	820	45.4	319	4	AAG72332
29	820	45.4	544	7	ADF70407
30	817	45.2	306	2	AAW35831
31	817	45.2	306	6	ABH58746
32	815	45.1	306	6	ABP57089
33	815	45.1	306	6	ABP81701
34	815	45.1	306	8	ADO05722
35	815	45.1	306	8	ADO29073
36	814.5	45.1	295	4	AAG72401
37	813.5	45.0	343	2	AAW59907
38	813.5	45.0	343	6	AAO19904
39	813.5	45.0	343	6	ABP81702
40	813.5	45.0	343	8	ADO29418
41	782	43.3	338	3	AAH18765
42	770	42.6	348	8	ADO29148
43	743	41.1	339	6	ABJ37891
44	743	41.1	348	3	AAH18764
45	743	41.1	348	5	ABH98168

ALIGNMENTS

RESULT 1					
ID	AAB49232	standard; protein; 339 AA.			
AC	AAB49232;				
DT	14-MAR-2001	(first entry)			
DE	Human SNORF3 receptor protein.				
KW	SNORF3; inflammation; arthritis; neurological disorder; infection;				
KW	bone disease; respiratory disorder; asthma; cancer; cardiovascular.				
OS	Homo sapiens.				
XX	WO200073449-A1.				
XX	07-DEC-2000.				
XX	26-MAY-2000; 2000WO-US014654.				
XX	28-MAY-1999; 99US-00322257.				
XX	06-OCT-1999; 99US-00413433.				
XX	(SYNA-) SYNAPTIC PHARM CORP.				
XX	Borowky BE, Ogoralek KL, Jones KA;				
XX	WPI; 2001-025252/03.				
XX	Nucleic acid encoding a mammalian (human, rat and mouse) SNORF3 receptor				
XX	which is useful for designing drugs for treating conditions such as a				
XX	chronic and acute inflammation, arthritis, neurological disorders and				
XX	microbial infections.				
XX	Claim 9; Fig 6; 227p; English.				
XX	The present invention relates to a mammalian SNORF3 receptor. SNORF3				
XX	antagonists and agonists are used to treat abnormalities brought about by				
XX	increased or decreased activity of the mammalian SNORF3 receptor. The				
XX	receptor is useful as a tool for designing drugs for treating conditions				
XX	such as a chronic and acute inflammation, arthritis, neurological				
XX	disorders, microbial infections, bone diseases, respiratory disorders				
XX	such as asthma, cancers, cardiovascular disorders				
XX	Sequence 339 AA;				
XX	Query Match	100.0%; Score 1806; DB 4; Length 339;			

Best Local Similarity 100.0%; Pred. No. 1,4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MMEPCNNINISCVKNMSNDVRAISYSLMWLIIITLVGNLIYVISISHPKQLHTPTNW 60
DB 1 MMEPCNNINISCVKNMSNDVRAISYSLMWLIIITLVGNLIYVISISHPKQLHTPTNW 60
QY 61 LIHSMATVDPLGLCLVMPYSMVRSAEHCWYFGEVFCIKHTSDIMLSASIFHLSPISID 120
DB 61 LIHSMATVDPLGLCLVMPYSMVRSAEHCWYFGEVFCIKHTSDIMLSASIFHLSPISID 120
QY 121 RYVAVCDPLRYKAKNNILVICWIFISMSVPAVFAFGMIFLEINFGAEIYYKVVHCRG 180
DB 121 RYVAVCDPLRYKAKNNILVICWIFISMSVPAVFAFGMIFLEINFGAEIYYKVVHCRG 180
QY 181 GCSVFSKISGLVLTFTSTFYIPGSIIMLCVYRIYLIANEQARLISDANOKLQIGLEMXNG 240
DB 181 GCSVFSKISGLVLTFTSTFYIPGSIIMLCVYRIYLIANEQARLISDANOKLQIGLEMXNG 240
QY 241 ISOSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
DB 241 ISOSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
QY 301 PMVYAFYFPMFRKALKMMLFGKIFQXDSRCKLFLLELSS 339
DB 301 PMVYAFYFPMFRKALKMMLFGKIFQXDSRCKLFLLELSS 339

RESULT 2
AA080973 standard; protein; 339 AA.
AC AAG80973;
AC AAG80973;
DT 28-AUG-2001 (first entry)
DE Human ngPCR56 #3.
KW G protein-coupled receptor; ngPCR; seven transmembrane receptor;
KW signal transduction; schizophrenia; thyroid disorder; renal failure;
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
KW cardiovascular disease; proliferative disorder; hormonal disorder;
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
KW attention deficit-hyperactivity disorder; attention deficit disorder;
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
KW neuroprotective.
OS Homo sapiens.
PN WO200136473-A2.
PD 25-MAY-2001.
PF 16-NOV-2000; 2000WO-US031581.
XX 16-NOV-1999; 99US-0165838P.
XX 17-NOV-1999; 99US-0166071P.
XX 19-NOV-1999; 99US-0166678P.
XX 28-DEC-1999; 99US-0173396P.
XX 22-FEB-2000; 2000US-0184129P.
XX 28-FEB-2000; 2000US-0185421P.
XX 28-FEB-2000; 2000US-0185549P.
XX 02-MAR-2000; 2000US-0186530P.
XX 03-MAR-2000; 2000US-0186811P.
XX 09-MAR-2000; 2000US-0188114P.
XX 17-MAR-2000; 2000US-0190310P.
XX 21-MAR-2000; 2000US-0190800P.
XX 20-APR-2000; 2000US-0198568P.
XX 02-MAY-2000; 2000US-0201190P.
XX 08-MAY-2000; 2000US-0203111P.
XX 25-MAY-2000; 2000US-0207094P.

```

(PHAA) PHARMACIA & UPJOHN CO.
PI Vogel I G, Wood IS, Parodi IA, Hiesch RR, Lind P, Slightow J,
PI Schellin KA, Kayes PS, Bannigan CM, Ruff V, Seijtitz T, Huff RM,
DR MPI: 2001-389826/41.
DR N-PSDB; AAH51013.

New G protein-coupled receptor (ngPCR-x) and its encoding polynucleotide
useful for diagnosing and treating e.g. schizophrenia.

Claim 37; Page 92; 261pp; English.

The present invention relates to novel G protein-coupled receptors
(ngPCR); where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27, 28,
31-38, 40, 41, 53-60 and their coding sequences. The present sequence is
one such G protein-coupled receptor. GPCRs are also known as seven
transmembrane receptors and function in signal transduction. The ngPCR
coding sequences are useful for screening a human to diagnose a disorder
affecting the brain or a genetic predisposition, specifically
schizophrenia. ngPCR are useful for identifying compounds useful for
treating schizophrenia. Detection of ngPCR in a sample is useful as a
diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
metabolic and cardiovascular diseases, proliferative disorders and
hormonal disorders. Modulators of ngPCR activity have the utility for
treating neurological disorders, including schizophrenia, ADHD/ADD
(attention deficit-hyperactivity disorder/attention deficit disorder),
CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
migraine and senile dementia. Additional disorders include inflammatory
conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
disorders, cancers, respiratory ailments such as asthma, and inflammatory
diseases e.g. inflammatory bowel disease

Sequence 339 AA;

Query Match 100.0%; Score 1806; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MMEPCNNINISCVKNMSNDVRAISYSLMWLIIITLVGNLIYVISISHPKQLHTPTNW 60
DB 1 MMEPCNNINISCVKNMSNDVRAISYSLMWLIIITLVGNLIYVISISHPKQLHTPTNW 60
QY 61 LIHSMATVDPLGLCLVMPYSMVRSAEHCWYFGEVFCIKHTSDIMLSASIFHLSPISID 120
DB 61 LIHSMATVDPLGLCLVMPYSMVRSAEHCWYFGEVFCIKHTSDIMLSASIFHLSPISID 120
QY 121 RYVAVCDPLRYKAKNNILVICWIFISMSVPAVFAFGMIFLEINFGAEIYYKVVHCRG 180
DB 121 RYVAVCDPLRYKAKNNILVICWIFISMSVPAVFAFGMIFLEINFGAEIYYKVVHCRG 180
QY 181 GCSVFSKISGLVLTFTSTFYIPGSIIMLCVYRIYLIANEQARLISDANOKLQIGLEMXNG 240
DB 181 GCSVFSKISGLVLTFTSTFYIPGSIIMLCVYRIYLIANEQARLISDANOKLQIGLEMXNG 240
QY 241 ISOSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
DB 241 ISOSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
QY 301 PMVYAFYFPMFRKALKMMLFGKIFQXDSRCKLFLLELSS 339
DB 301 PMVYAFYFPMFRKALKMMLFGKIFQXDSRCKLFLLELSS 339

RESULT 3
ABJ04073 standard; protein; 339 AA.
AC ABJ04073;
AC ABJ04073;
DT 11-OCT-2002 (first entry)
XX

```


DE Human G protein coupled receptor hRUP33.
XX Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
KW hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37.
XX Homo sapiens.
XX WO200242461-A2.
XX
XX 30-MAY-2002.
XX
XX 26-NOV-2001; 2001WO-US044386.
XX
XX 27-NOV-2000; 2000US-0253404P.
XX 12-DEC-2000; 2000US-0255366P.
XX 20-FEB-2001; 2001US-0270266P.
XX 20-FEB-2001; 2001US-0270286P.
XX 06-APR-2001; 2001US-0282032P.
XX 06-APR-2001; 2001US-0282356P.
XX 06-APR-2001; 2001US-0282358P.
XX 06-APR-2001; 2001US-0282365P.
XX 14-MAY-2001; 2001US-0290917P.
XX 31-JUL-2001; 2001US-0309208P.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX WPI; 2002-566565/60.
XX DR N-PSDB; ABT04871.
XX
XX Novel endogenous and non-endogenous versions of G protein-coupled
PT receptor useful for identification of candidate compounds as receptor
PT agonists or antagonists for use as therapeutic agents.
XX
XX Claim 21; Page 66-67; 84pp; English.
XX
XX The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a GPCR protein of the invention
XX
XX Sequence 339 AA;
SQ
Query Match 100.0%; Score 1806; DB 5; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
ABG93791
ID ABG93791 standard; protein; 339 AA.
XX
XX ABG93791;
XX
XX 26-NOV-2002 (first entry)
XX
XX Human G protein-coupled receptor protein, nGPR-56, #2.
XX
XX Human; receptor; G protein-coupled receptor; GPCR; nGPR; beGPCR;
KW nG protein coupled receptor; communication; serpentine structure;
KW seven transmembrane receptor; 7TM; mental disorder; diagnosis;
KW genetic predisposition; brain; immune response; gene therapy;
KW anxiety disorder; depression; bipolar disorder; schizophrenia;
KW Huntington's disease; dyskinesia; manic depression; stroke;
KW Parkinson's disease; Alzheimer's disease; diabetes; inflammation; wound;
KW tranquiliser.
XX
XX Homo sapiens.
XX
XX WO200264789-A1.
XX
XX 22-AUG-2002. ✓
XX
XX 14-FEB-2001; 2001WO-US004641.
XX
XX 14-FEB-2001; 2001WO-US004641.
XX
XX (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Lind P, Parodi LA, Vogel I, Wood LS;
XX WPI; 2002-674879/72.
XX DR N-PSDB; ABS70246.
XX
XX New nucleic acids and polypeptides of the nG protein-coupled receptor,
PT useful for treating or diagnosing a mental disorder or a disorder
PT affecting the brain, e.g. anxiety disorders, schizophrenia, stroke or
PT Parkinson's disease.
XX
XX Example 1; Page 87; 244pp; English.
XX
XX The invention discloses an isolated human polypeptide, and encoding
CC nucleic acid, for a G protein-coupled receptor (GPCR), particularly the
CC nG protein coupled receptor-14 (nGPR-14). GPCRs are vital in the
CC communication between cells and their environment and are characterised
CC by a serpentine structure that passes through the cell membrane seven
CC times, hence the reason such receptors are sometimes called seven
CC transmembrane receptors (7TM). The polynucleotides and polypeptides are
CC useful for identifying an nGPR allelic variant that correlates with a
CC mental disorder, for isolating an antibody that binds to an epitope of
CC the polypeptide, for identifying a compound that binds the polypeptide or
CC polynucleotide and/or modulates its biological activity, for screening a
CC human subject to diagnose a disorder, or a genetic predisposition to a
CC disorder, affecting the brain or a genetic disposition to the disorder,
CC for identifying compounds useful for the treatment of a mental disorder,
CC and for identifying a compound useful as a modulator of binding between
CC nGPR-14 and a binding partner of nGPR-14. The polypeptide is also
CC useful for inducing an immune response in a mammal. The nucleic acid or
CC polypeptide is particularly useful, using gene therapy, for treating e.g.
CC anxiety disorders, depression, bipolar disorder, schizophrenia,
CC Huntington's disease, dyskinesias, manic depression, stroke, Parkinson's
CC disease or Alzheimer's disease. The nucleic acid and polypeptide may also
CC be used for treating diabetes, inflammation or wounds. The sequences
CC presented in ab beGPCRs) proteins
XX referred to as beGPCRs) proteins
XX
XX Sequence 339 AA;
SQ
Query Match 100.0%; Score 1806; DB 5; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;

Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPEPCNNINISCVKNNMSNDVRASLYSLMWLIIITLVGNLVIYVISHFKQLHPTNM 60

Db 1 MPEPCNNINISCVKNNMSNDVRASLYSLMWLIIITLVGNLVIYVISHFKQLHPTNM 60

QY 61 LIHSMATVDLGLGLVMPYSNVRSAEHCWYFGEVCKIHSTIDIMLSASIFHLSFISID 120

Db 61 LIHSMATVDLGLGLVMPYSNVRSAEHCWYFGEVCKIHSTIDIMLSASIFHLSFISID 120

QY 121 RYAVACDPLRYKAKNNILVICMIFISWSVPAVAFGMIFLEINFKGAEIYYKHVHCRG 180

Db 121 RYAVACDPLRYKAKNNILVICMIFISWSVPAVAFGMIFLEINFKGAEIYYKHVHCRG 180

QY 181 GCSVFSKISGVLTFMTSFYIPGSIMLCVYRYRYLAKEOARLISDANKLOIGLEMKNG 240

Db 181 GCSVFSKISGVLTFMTSFYIPGSIMLCVYRYRYLAKEOARLISDANKLOIGLEMKNG 240

QY 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTN 300

Db 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTN 300

QY 301 PMVYAFYPMFRKALKMFLGKIFOKDSRCKLFLFLSS 339

Db 301 PMVYAFYPMFRKALKMFLGKIFOKDSRCKLFLFLSS 339

RESULT 5
ADH48916
ID ADH48916 standard; protein; 339 AA.
XX
XX ADH48916;
DT 25-MAR-2004 (first entry)
DE NOV84B protein sequence, SEQ ID 200.
XX
XX Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;
KW hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV84B;
KW GPCR-like protein; chromosome 6.
XX
XX Homo sapiens.
OS
PN WO200268652-A2.
PD 06-SEP-2002. ✓
PF 26-FEB-2002; 2002WO-US005910.
XX
XX 26-FEB-2001; 2001US-0271646P.
PR 27-FEB-2001; 2001US-0271840P.
PR 28-FEB-2001; 2001US-0272404P.
PR 28-FEB-2001; 2001US-0272405P.
PR 28-FEB-2001; 2001US-0272410P.
PR 28-FEB-2001; 2001US-0272414P.
PR 02-MAR-2001; 2001US-0272787P.
PR 02-MAR-2001; 2001US-0272922P.
PR 02-MAR-2001; 2001US-0273048P.
PR 02-MAR-2001; 2001US-0273300P.
PR 16-MAR-2001; 2001US-0276401P.
PR 20-MAR-2001; 2001US-0278660P.
PR 30-MAR-2001; 2001US-0280039P.
PR 30-MAR-2001; 2001US-0280234P.
PR 02-APR-2001; 2001US-0280818P.
PR 12-APR-2001; 2001US-0283443P.
PR 23-APR-2001; 2001US-0285754P.
PR 24-APR-2001; 2001US-0286036P.
PR 03-MAY-2001; 2001US-0288353P.
PR 17-MAY-2001; 2001US-0291703P.
PR 31-MAY-2001; 2001US-0294834P.
PR 20-JUN-2001; 2001US-0296959P.
PR 21-JUN-2001; 2001US-0299845P.

PR 05-JUL-2001; 2001US-0303242P.
PR 13-AUG-2001; 2001US-0311981P.
PR 16-AUG-2001; 2001US-0312858P.
PR 17-AUG-2001; 2001US-0313280P.
PR 29-AUG-2001; 2001US-0315614P.
PR 17-SEP-2001; 2001US-0322818P.
PR 25-FEB-2002; 2002US-00322818.
XX
XX (CUBA-) CUBAGEN CORP.
XX
XX Aisbrook JP, Anderson DW, Ballinger RA, Boidog FL, Burgess CE;
PI Casman SJ, Silleman KE, Gangoli EA, Gerlach VL, Gilbert JA,
PI Gorman L, Guo X, Guzey VY, Kekuda R, Li L, Liu X, Malpankar UM,
PI Miller CE, Millet I, Padigaru M, Patcurajan M, Pena CE, Peyman JA,
PI Rastelli L, Shenoy SG, Shimkets RA, Smithson G, Szytek KA, Stone DJ,
PI Taupier RJ, Tchernev VT, Vernet CM, Zernušen BD;
XX
XX WPI; 2002-698672/75.
XX
XX N-PsDB; ADH48915.
XX
XX PT New NOVX polypeptides or polynucleotides, useful for preventing or
PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
PT obesity or cancer.
XX
XX PS Claim 1; Page 452; 923pp; English.

XX
XX The present invention relates to novel human NOVX proteins, where X is
XX any number from 1 to 91 and their coding sequences. The proteins and
XX coding sequences are useful for preventing or treating disorders or
XX syndromes e.g., atherosclerosis, hypertension, obesity or cancer. NOV73-
XX NOV91 are GPCR-like proteins. The coding sequence for the present NOV
XX maps to chromosome 6.

XX
XX Sequence 339 AA:

Query Match 100.0%; Score 1806; DB 5; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPEPCNNINISCVKNNMSNDVRASLYSLMWLIIITLVGNLVIYVISHFKQLHPTNM 60

Db 1 MPEPCNNINISCVKNNMSNDVRASLYSLMWLIIITLVGNLVIYVISHFKQLHPTNM 60

QY 61 LIHSMATVDLGLGLVMPYSNVRSAEHCWYFGEVCKIHSTIDIMLSASIFHLSFISID 120

Db 61 LIHSMATVDLGLGLVMPYSNVRSAEHCWYFGEVCKIHSTIDIMLSASIFHLSFISID 120

QY 121 RYAVACDPLRYKAKNNILVICMIFISWSVPAVAFGMIFLEINFKGAEIYYKHVHCRG 180

Db 121 RYAVACDPLRYKAKNNILVICMIFISWSVPAVAFGMIFLEINFKGAEIYYKHVHCRG 180

QY 181 GCSVFSKISGVLTFMTSFYIPGSIMLCVYRYRYLAKEOARLISDANKLOIGLEMKNG 240

Db 181 GCSVFSKISGVLTFMTSFYIPGSIMLCVYRYRYLAKEOARLISDANKLOIGLEMKNG 240

QY 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTN 300

Db 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTN 300

QY 301 PMVYAFYPMFRKALKMFLGKIFOKDSRCKLFLFLSS 339

Db 301 PMVYAFYPMFRKALKMFLGKIFOKDSRCKLFLFLSS 339

RESULT 6
ADH48908
ID ADH48908 standard; protein; 339 AA.
XX
XX ADH48908;
AC
AC ADH48908;
DT 25-MAR-2004 (first entry)
XX
XX NOV81B protein sequence, SEQ ID 192.

XX Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;
KW hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV61B;
KW GPCR-like protein; chromosome 6.
XX Homo sapiens.
XX WO200268652-A2.
XX 06-SEP-2002. /
XX
XX 26-FEB-2002; 2002WO-US005910.
XX
XX 26-FEB-2001; 2001US-0271646P.
XX 27-FEB-2001; 2001US-0271840P.
XX 28-FEB-2001; 2001US-0272404P.
XX 28-FEB-2001; 2001US-0272405P.
XX 28-FEB-2001; 2001US-0272410P.
XX 28-FEB-2001; 2001US-0272414P.
XX 02-MAR-2001; 2001US-0272787P.
XX 02-MAR-2001; 2001US-0272922P.
XX 02-MAR-2001; 2001US-0273048P.
XX 02-MAR-2001; 2001US-0273300P.
XX 16-MAR-2001; 2001US-0276401P.
XX 20-MAR-2001; 2001US-0277324P.
XX 20-MAR-2001; 2001US-0278660P.
XX 30-MAR-2001; 2001US-0280039P.
XX 30-MAR-2001; 2001US-0280234P.
XX 02-APR-2001; 2001US-0280818P.
XX 12-APR-2001; 2001US-0283443P.
XX 23-APR-2001; 2001US-0285754P.
XX 24-APR-2001; 2001US-0286096P.
XX 03-MAY-2001; 2001US-0288353P.
XX 17-MAY-2001; 2001US-0291703P.
XX 31-MAY-2001; 2001US-0294834P.
XX 20-JUN-2001; 2001US-0299695P.
XX 21-JUN-2001; 2001US-0299845P.
XX 05-JUL-2001; 2001US-0303242P.
XX 13-AUG-2001; 2001US-0311981P.
XX 16-AUG-2001; 2001US-0312858P.
XX 17-AUG-2001; 2001US-0313280P.
XX 29-AUG-2001; 2001US-0315614P.
XX 17-SEP-2001; 2001US-0322818P.
XX 25-FEB-2002; 2002US-00322818.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Alsebrook JP, Anderson DM, Ballinger RA, Boldog FL, Burgess CE;
PI Caeman SJ, Ellerman KE, Gangolli EA, Gerlach VL, Gilbert JA;
PI Gorman L, Guo X, Gusev VY, Kekuda R, Li L, Liu X, Malysankar UM;
PI Miller CE, Muller I, Padigaru M, Patturajan M, Pena CA, Peyman JA;
PI Rastelli L, Shenoy SG, Shinkens RA, Smithson G, Spytek KA, Stone DJ;
PI Taupier RJ, Tchernev VT, Vernet CM, Zernusen BJ.
XX WPI; 2002-698672/75.
XX N-PSDB; ADH48907.
XX
XX New NOVX polypeptides or polynucleotides, useful for preventing or
PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
PT obesity or cancer.
XX
XX Claim 1, Page 440; 923pp; English.
XX
XX The present invention relates to novel human NOVX proteins, where X is
CC any number from 1 to 91 and their coding sequences. The proteins and
CC coding sequences are useful for preventing or treating disorders or
CC syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOV73-
CC NOV91 are GPCR-like proteins. The coding sequence for the present NOV
XX maps to chromosome 6.
XX
XX Sequence 339 AA;
SQ

Query Match 100.0%; Score 1806; DB 5; Length 339;

Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMPFGHNINISCVKNMNSNDVRASLYSLMTLITTLVGNLIVVSIHSFQHLPTNW 60
DB 1 MMPFGHNINISCVKNMNSNDVRASLYSLMTLITTLVGNLIVVSIHSFQHLPTNW 60
QY 61 LIHSMATVDFLLGCLVMPYSWVSAHCHMVEGVECKIHTSTDIMSSASIFHLSPISD 120
DB 61 LIHSMATVDFLLGCLVMPYSWVSAHCHMVEGVECKIHTSTDIMSSASIFHLSPISD 120
QY 121 RYVAVCDPLRYKAKNNIIVICMIFISWSVPAVAFGMI FLELNFKGAEIYYKVVHCRG 180
DB 121 RYVAVCDPLRYKAKNNIIVICMIFISWSVPAVAFGMI FLELNFKGAEIYYKVVHCRG 180
QY 181 GCSVFEFSKISGVLTFMTSPFI PGSI MLGVYRIYLAEOQRLISDANOKLQIGLEMKNG 240
DB 181 GCSVFEFSKISGVLTFMTSPFI PGSI MLGVYRIYLAEOQRLISDANOKLQIGLEMKNG 240
QY 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
DB 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFN 300
QY 301 PMVYAFYPWPFRKALKMLFGKIFQKDSRCKLFLIELSS 339
DB 301 PMVYAFYPWPFRKALKMLFGKIFQKDSRCKLFLIELSS 339
RESULT 7
ADH48906
ID ADH48906 standard; protein; 339 AA.
XX
XX ADH48906;
XX
XX 25-MAR-2004 (first entry)
XX
XX NOV61A protein sequence, SEQ ID 190.
XX
XX Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;
KW hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV61A;
KW GPCR-like protein; chromosome 6.
XX
XX Homo sapiens.
XX
XX WO200268652-A2.
XX PN
XX 06-SEP-2002. /
XX
XX 26-FEB-2002; 2002WO-US005910.
XX PF
XX
XX 26-FEB-2001; 2001US-0271646P.
XX 27-FEB-2001; 2001US-0271840P.
XX 28-FEB-2001; 2001US-0272404P.
XX 28-FEB-2001; 2001US-0272405P.
XX 28-FEB-2001; 2001US-0272410P.
XX 28-FEB-2001; 2001US-0272414P.
XX 02-MAR-2001; 2001US-0272787P.
XX 02-MAR-2001; 2001US-0272922P.
XX 02-MAR-2001; 2001US-0273048P.
XX 02-MAR-2001; 2001US-0273300P.
XX 16-MAR-2001; 2001US-0276401P.
XX 20-MAR-2001; 2001US-0277324P.
XX 20-MAR-2001; 2001US-0278660P.
XX 30-MAR-2001; 2001US-0280039P.
XX 30-MAR-2001; 2001US-0280234P.
XX 02-APR-2001; 2001US-0280818P.
XX 12-APR-2001; 2001US-0283443P.
XX 23-APR-2001; 2001US-0285754P.
XX 24-APR-2001; 2001US-0286096P.
XX 03-MAY-2001; 2001US-0288353P.
XX 17-MAY-2001; 2001US-0291703P.
XX 31-MAY-2001; 2001US-0294834P.
XX 20-JUN-2001; 2001US-0299695P.

PR 21-JUN-2001; 2001US-0299845P.
 PR 05-JUL-2001; 2001US-0303242P.
 PR 13-AUG-2001; 2001US-0311981P.
 PR 16-AUG-2001; 2001US-0312858P.
 PR 17-AUG-2001; 2001US-0313280P.
 PR 29-AUG-2001; 2001US-0315614P.
 PR 17-SEP-2001; 2001US-0322818P.
 PR 25-FEB-2002; 2002US-00322818.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alsobrook JP, Anderson DW, Ballinger RA, Boldog FL, Burgess CE,
 PI Gorman SJ, Ellerman KE, Gangoli EA, Gerlach VL, Gilbert JA,
 PI Gorman LJ, Guo X, Gusev VY, Kikuda R, Li L, Liu X, Malvankar UM,
 PI Miller CE, Millet I, Padigar M, Paturajan M, Pena CE, Peyman JA,
 PI Rastelli L, Shenoy SG, Shinkens RA, Smithson G, Spytek KA, Stone DJ,
 PI Taupier RJ, Tchernev VT, Vernet CM, Zernusen BD;
 PI
 DR WPI; 2002-698672/75.
 DR N-PSDB; ADH48905.
 XX
 PT New NOVX polypeptides or polynucleotides, useful for preventing or
 PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
 PT obesity or cancer.
 PS
 PS Claim 1; Page 439; 923pp; English.
 XX
 CC The present invention relates to novel human NOVX proteins, where X is
 CC any number from 1 to 91 and their coding sequences. The proteins and
 CC coding sequences are useful for preventing or treating disorders or
 CC syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOVX-
 CC NOV91 are GPCR-like proteins. The coding sequence for the present NOV
 CC maps to chromosome 6.
 XX
 SQ Sequence 339 AA;
 Query Match 100.0%; Score 1806; DB 5; Length 339;
 Best Local Similarity 100.0%; Pred. No. 1.4e-196;
 Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MMPECHNINISCVKNNMSNDVRASLYSLMVLIIITLWGNLIYVSI SHFQQLHTPTNW 60
 Db 1 MMPECHNINISCVKNNMSNDVRASLYSLMVLIIITLWGNLIYVSI SHFQQLHTPTNW 60
 QY 61 LIHSAATVDFLLGCLVMPYSWVSAEHQWYEGVEVCKIHTSDIMLSASIFHLISFISID 120
 Db 61 LIHSAATVDFLLGCLVMPYSWVSAEHQWYEGVEVCKIHTSDIMLSASIFHLISFISID 120
 QY 121 RYAAVCDPLRYAKNNIIIVICMIFISVPAVFAFGMIFELNKGABEIIYKXVHCRG 180
 Db 121 RYAAVCDPLRYAKNNIIIVICMIFISVPAVFAFGMIFELNKGABEIIYKXVHCRG 180
 QY 121 RYAAVCDPLRYAKNNIIIVICMIFISVPAVFAFGMIFELNKGABEIIYKXVHCRG 180
 Db 121 RYAAVCDPLRYAKNNIIIVICMIFISVPAVFAFGMIFELNKGABEIIYKXVHCRG 180
 QY 181 GCSVFPKISGVLTFTMTSFYIPGSIIMLCVYRIYLIKAEQARLISDAOKIQIGEMKNG 240
 Db 181 GCSVFPKISGVLTFTMTSFYIPGSIIMLCVYRIYLIKAEQARLISDAOKIQIGEMKNG 240
 QY 241 ISQSKERAAVKTLLGIVMGVFLICWCPFTCTVMDPFLHYIIPPLNDVLIIFGYLNTSFN 300
 Db 241 ISQSKERAAVKTLLGIVMGVFLICWCPFTCTVMDPFLHYIIPPLNDVLIIFGYLNTSFN 300
 QY 301 PMVYAFYFPMFPRKALKMMFLGKIPKDKSSRCKLPLEISS 339
 Db 301 PMVYAFYFPMFPRKALKMMFLGKIPKDKSSRCKLPLEISS 339
 RESULT 8
 ABP81732
 ID ABP81732 standard; protein; 339 AA.
 AC ABP81732;
 DT 04-MAR-2003 (first entry)

DE Human trace amine receptor 1 protein SEQ ID NO:639.
 XX
 XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KW G protein-coupled receptor modulator; antibody; immune-related disease;
 KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KW immunological-related cell proliferative disease; autoimmune disease;
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; memory loss;
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KW ulcer.
 XX
 XX Homo sapiens.
 OS
 PN WO200261087-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 19-DEC-2001; 2001WO-US050107.
 XX
 PR 19-DEC-2000; 2000US-0257144P.
 XX
 PA (LIFE-) LIFESPAN BIOSCIENCES INC.
 XX
 PI Burner GC, Roush CL, Brown JP;
 PI
 DR WPI; 2003-046718/04.
 DR N-PSDB; ABZ42578.
 XX
 PT New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 PS
 PS Disclosure; Fig 1; 523pp; English.
 XX
 CC The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnoses. ABZ42523 to ABZ42869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 339 AA;
 Query Match 100.0%; Score 1806; DB 6; Length 339;
 Best Local Similarity 100.0%; Pred. No. 1.4e-196;
 Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MMPECHNINISCVKNNMSNDVRASLYSLMVLIIITLWGNLIYVSI SHFQQLHTPTNW 60
 Db 1 MMPECHNINISCVKNNMSNDVRASLYSLMVLIIITLWGNLIYVSI SHFQQLHTPTNW 60

QY 61 LIHSMATVDFLLGCLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLFSISD 120
CC
CC LIHSMATVDFLLGCLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLFSISD 120
Db 61 LIHSMATVDFLLGCLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLFSISD 120
QY 121 RYAVACDPLRYAKANNILVICMIFISWSPVAFAGMIFLELNFKAEBIYYKKVHCRC 180
CC
CC RYAVACDPLRYAKANNILVICMIFISWSPVAFAGMIFLELNFKAEBIYYKKVHCRC 180
Db 121 RYAVACDPLRYAKANNILVICMIFISWSPVAFAGMIFLELNFKAEBIYYKKVHCRC 180
QY 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYLIAEQARLISDANQKQIGLEMKNG 240
CC
CC 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYLIAEQARLISDANQKQIGLEMKNG 240
Db 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYLIAEQARLISDANQKQIGLEMKNG 240
QY 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIPTLNDVLIMFGYLNSTFN 300
CC
CC 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIPTLNDVLIMFGYLNSTFN 300
Db 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIPTLNDVLIMFGYLNSTFN 300
QY 301 PMVYAFYPPWPRKALKMMLFGKIFOKDSRCKLFLLELS 339
CC
CC 301 PMVYAFYPPWPRKALKMMLFGKIFOKDSRCKLFLLELS 339
Db 301 PMVYAFYPPWPRKALKMMLFGKIFOKDSRCKLFLLELS 339
RESULT 9
ADO29662
ID ADO29662 standard; protein; 339 AA.
AC ADO29662;
XX
XX 29-JUL-2004 (first entry)
DT
DE Human GPCR TARI, SEQ ID NO:764.
XX
KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
KW cytostatic; antineoplastic; vasodilator; antidiabetic;
KW CNS; central nervous system; respiratory; antidiabetic; antidiabetic;
KW vitruclide; hepatotropic; antibacterial; antianemic; antidiabetic;
KW dermatological; anticancer; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotoxic; gene therapy; GPCR modulator; human;
KW receptor.
XX
OS Homo sapiens.
XX
PN WO2004040000-A2.
XX
PD 13-MAY-2004. /
XX
PF 09-SEP-2003; 2003WO-US028226.
XX
PR 09-SEP-2002; 2003US-0409303P.
XX
PR 09-APR-2003; 2003US-0461329P.
XX
PA (PRIM-) PRIMAL INC.
XX
PI Galanarlis GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
XX
PI Madisen L, Meliwein KL, Pavlova MN, Vasiliadis D, Zeng H;
XX
DR WPI; 2004-390329/36.
XX
DR N-PSDB; ADO30050.
XX
PT Novel mammalian G protein coupled receptors, useful for identifying
XX
PT compounds that modulates diagnosing and treating disease condition
XX
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
XX
PT pectoris, Parkinson's disease.
XX
PS Claim 151; SEQ ID NO 764; 542bp; English.

XX The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention, methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention, methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridize to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR of the
CC invention. Note: The full sequence data for this patent did not form part
CC of the printed specification; those sequences not shown were obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences.
XX

Seq Sequence 339 AA:

Query Match 100.0%; Score 1806; DB 8; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMPFCHNINISCVKNNMSNDVRAALYSIMLIIITLVGNLIYVVISHFQQLHTPTNW 60
Db 1 MMPFCHNINISCVKNNMSNDVRAALYSIMLIIITLVGNLIYVVISHFQQLHTPTNW 60
QY 61 LIHSMATVDFLLGCLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLFSISD 120
CC
CC 61 LIHSMATVDFLLGCLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLFSISD 120
Db 61 LIHSMATVDFLLGCLVMPYSAHCHWYFGEVFCCKIHTSTDIMSSASIFHLFSISD 120
QY 121 RYAVACDPLRYAKANNILVICMIFISWSPVAFAGMIFLELNFKAEBIYYKKVHCRC 180
CC
CC 121 RYAVACDPLRYAKANNILVICMIFISWSPVAFAGMIFLELNFKAEBIYYKKVHCRC 180
Db 121 RYAVACDPLRYAKANNILVICMIFISWSPVAFAGMIFLELNFKAEBIYYKKVHCRC 180
QY 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYLIAEQARLISDANQKQIGLEMKNG 240
CC
CC 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYLIAEQARLISDANQKQIGLEMKNG 240
Db 181 GCSVFFSKISGVLTFMTSFYIPGSIMLCVYRYIYLIAEQARLISDANQKQIGLEMKNG 240
QY 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIPTLNDVLIMFGYLNSTFN 300
CC
CC 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIPTLNDVLIMFGYLNSTFN 300
Db 241 ISOSKERAQVKTGLIGVGVFLICWCPFFICTYMDPFLHIPTLNDVLIMFGYLNSTFN 300
QY 301 PMVYAFYPPWPRKALKMMLFGKIFOKDSRCKLFLLELS 339
CC
CC 301 PMVYAFYPPWPRKALKMMLFGKIFOKDSRCKLFLLELS 339
Db 301 PMVYAFYPPWPRKALKMMLFGKIFOKDSRCKLFLLELS 339
RESULT 10
ADS73691
ID ADS73691 standard; protein; 339 AA.
XX
XX ADS73691;
AC ADS73691;
XX
XX 16-DEC-2004 (first entry)
DT
XX
DE Human trace amine receptor 1 (TAR) polypeptide.

XX TAL; GPCR; trace amine receptor 1; antianaemic; antiinflammatory;
KW cardiovascular; neuroprotective; respiratory; uropathic;
KW G protein coupled receptor; gene therapy; human; CSRL.
OS Homo sapiens.
FN WO2004083851-A2.
PN 30-SEP-2004.
PD 04-MAR-2004; 2004MO-EP002182.
PF 18-MAR-2003; 2003EP-00006101.
PR (PARB) BAYER HEALTHCARE AG.
PA Golz S, Brueggemeier U, Geerts A;
PI WPI; 2004-691179/67.
XX N-PSDB; ADST3690.
DR Screening for therapeutic agents, useful for treating e.g.,
XX cardiovascular and inflammatory disorders, comprising contacting a test
PT compound with trace amine receptor 1 (TAL) polypeptide and detecting
PT their binding.
PS Disclosure; SEQ ID NO 2; 117BP; English.
XX
XX The invention relates to screening for therapeutic agents and involves
CC contacting a test compound with trace amine receptor 1 (TAL) polypeptide
CC or polynucleotide and detecting binding of the test compound to TAL
CC polypeptide or polynucleotide, or determining TAL polypeptide activity at
CC a certain test compound concentration or in the absence of the test
CC compound and at a different concentration of the test compound. The
CC therapeutic agents are useful in treating disease such as cardiovascular
CC disorders, inflammatory disorders, hematological disorders, respiratory
CC disorders, neurological disorders or urological disorders in a mammal. The
CC regulators of TAL are useful in the preparation of a pharmaceutical
CC composition for treating the above diseases and also useful for the
CC regulation of TAL activity in a mammal having the disease. The nucleotide
CC sequences encoding TAL are useful as hybridization probes, in
CC constructing oligomers for PCR, for chromosome and gene mapping, in the
CC recombinant production of TAL, in generating antisense DNA or RNA and in
CC molecular biology techniques that have not yet been developed. TAL are
CC useful for immunizing a mammal to produce polyclonal antibodies and for
CC diagnostic purposes. The present sequence represents a human TAL
CC receptor, a seven transmembrane G protein coupled receptor polypeptide.
XX
SQ Sequence 339 AA;
Query Match 100.0%; Score 1806; DB 8; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,48-196;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 ISQSKERKAVKTLGIWGVFLICWCPFLICTVMDPFLHYIIPTLNDVLIWGYLSTFN 300
QY 301 PMVYAFPPWPRKALKMLFGKIFOKDSRRCLFLELSS 339
DB 301 PMVYAFPPWPRKALKMLFGKIFOKDSRRCLFLELSS 339
RESULT 11
AAB82970
ID AAB82970 standard; protein, 338 AA.
XX AAB82970;
XX 21-DEC-2001 (first entry)
XX Human G protein coupled receptor.
XX G protein coupled receptor; GPCR; aminergic receptor;
KW signal transduction; therapy; human.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Modified-site 9..12 /note= "Asn is N-glycosylated"
FT Modified-site 16..19 /note= "Asn is N-glycosylated"
FT Domain 27..47 /note= "Asn is N-glycosylated"
FT Domain 61..81 /note= "transmembrane domain 1"
FT Domain 82..89 /note= "transmembrane domain 2"
FT Modified-site 99..119 /note= "Tyr is O-phosphorylated by tyrosine kinase"
FT Domain 99..102 /note= "transmembrane domain 3"
FT Modified-site 108..124 /note= "Thr is O-phosphorylated by casein kinase II"
FT Peptide 136..156 /note= "G protein coupled receptor signature"
FT Domain 165..172 /note= "transmembrane domain 4"
FT Modified-site 189..209 /note= "Tyr is O-phosphorylated by tyrosine kinase"
FT Domain 239..244 /note= "transmembrane domain 5"
FT Modified-site 252..272 /note= "N-myristoylated"
FT Domain 253..258 /note= "transmembrane domain 6"
FT Modified-site 270..273 /note= "N-myristoylated"
FT Modified-site 283..286 /note= "Thr is O-phosphorylated by casein kinase II"
FT Domain 287..307 /note= "Thr is O-phosphorylated by casein kinase II"
FT Modified-site 295..298 /note= "transmembrane domain 7"
FT Modified-site 327..329 /note= "Asn is N-glycosylated"
FT /note= "Ser is O-phosphorylated by protein kinase C"
XX
XX WO200172841-A2.
XX 04-OCT-2001.
XX 27-MAR-2001; 2001MO-US009660.
XX 27-MAR-2000; 2000US-0192311P.
PR 04-AUG-2000; 2000US-00633145.
XX
XX (PEKE) PE CORP NY.

XX 18-DEC-2003 (first entry)
 XX Human GPCR protein, SEQ ID NO 124.
 DE
 XX G protein-coupled receptor; GPCR; antibacterial; fungicide; protozoacide;
 KM virocidic; antirheumatic; antiarthritic; tranquiliser; antidiabetic;
 KM osteopathic; neurotropic; neuroprotective; anorectic; cardiant;
 KM neuroleptic; cytostatic; antiparkinsonian; hypotensive; hypertensive;
 KM antitumor; antiallergic; anticonvulsant; analgesic; infection; COPD;
 KM rheumatoid arthritis; chronic obstructive pulmonary diseases; COPD;
 KM asthma; non-insulin dependent diabetes; obesity; osteoporosis;
 KM Alzheimer's disease; age-related macular degeneration;
 KM myocardial infarction; schizophrenia; osteoarthritis; cancer;
 KM Parkinson's disease; congestive heart failure; hypertension; hypertension;
 KM ulcer; allergy; benign prostatic hyperplasia; seizure disorder; anxiety;
 KM obsessive compulsive disorder; Cushing's syndrome; hypopituitarism; pain;
 KM human.
 XX Homo sapiens.
 OS
 XX WO200300893-A2.
 XX
 XX 03-JAN-2003.
 PD
 XX 24-JUN-2002; 2002WO-IB002357.
 PF
 XX 26-JUN-2001; 2001US-0301095P.
 PR 06-NOV-2001; 2001US-0333185P.
 XX
 XX (DECO-) DECODE GENETICS EHF.
 PA
 XX Martinez RMA, Sigurdsson GR;
 PI
 XX WPT, 2003-210155/20.
 DR N-PSDB; ADC12791.
 DR
 XX
 PT New G protein-coupled receptor (GPCR) genes and polypeptide, useful for
 PT diagnosing diseases associated with a GPCR, or in gene therapy for
 PT treating e.g. obesity, osteoporosis, Alzheimer's, cancers or congestive
 PT heart failure.
 PT
 XX
 PS Claim 10; SEQ ID NO 124; 253pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid of a G protein-
 CC coupled receptor (GPCR) gene comprising any of 62 sequences of 912-2454
 CC bp, or its complements; a GPCR polypeptide comprising any of 62 sequences
 CC of 291-818 amino acids; or a nucleic acid that hybridises, under high
 CC stringency conditions, with any of the 62 GPCR sequences or any of their
 CC complements. The GPCR agents of the invention have the following
 CC activities: antibacterial, fungicide, protozoacide, virucide,
 CC antirheumatic, tranquiliser, antiarthritic, antidiabetic, osteopathic,
 CC neurotropic, neuroprotective, anorectic, cardiant, neuroleptic, cytostatic,
 CC antiparkinsonian, hypotensive, hypertensive, antitumor, antiallergic,
 CC anticonvulsant, and analgesic. The GPCR therapeutic agent, particularly a
 CC GPCR gene agonist or antagonist, is useful for treating a disease or
 CC condition associated with a GPCR in an individual. The nucleic acid cited
 CC above, which is 100 or fewer nucleotides in length, is useful for
 CC assaying a sample for the presence of the GPCR gene nucleic acid or a
 CC GPCR gene nucleic acid with at least one nucleotide difference from a
 CC first nucleic acid, or for diagnosing a susceptibility to a disease or
 CC conditions associated with a GPCR. These diseases include infections
 CC (e.g. bacterial, fungal, protozoan or viral), rheumatoid arthritis,
 CC chronic obstructive pulmonary diseases (COPD), asthma, non-insulin
 CC dependent diabetes, obesity, osteoporosis, Alzheimer's disease, age-
 CC related macular degeneration, myocardial infarction, schizophrenia,
 CC osteoarthritis, cancers, Parkinson's diseases, congestive heart failure,
 CC hypertension, hyperextension, ulcers, allergies, benign prostatic
 CC hyperplasia, seizure disorder, anxiety, obsessive compulsive disorder,
 CC Cushing's syndrome, hypopituitarism, or pain. This sequence represents
 CC one of the 62 GPCR proteins of the invention.
 CC
 XX Sequence 311 AA;
 SQ

Query Match 85.2%; Score 1539; DB 7; Length 311;
 Best Local Similarity 100.0%; Pred. No. 3,3e-166;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 29 LMVLIITLTVGNLIVIVISHPKQLHTPTNWLHSMATVDFLLGLVMPYSMAASHC 88
 DB 23 LMVLIITLTVGNLIVIVISHPKQLHTPTNWLHSMATVDFLLGLVMPYSMAASHC 82
 QY 89 MYFGVFEKIKTSDIMLSSASIFLSTISIDRYAVCDPLRYKKNLIVCVNIFLSW 148
 DB 83 MYFGVFEKIKTSDIMLSSASIFLSTISIDRYAVCDPLRYKKNLIVCVNIFLSW 142
 QY 149 SVPNFAFGMTFLEINFGAEIYYKHVCHGCGCVFESKISGVLTFMTSPYISIMLC 208
 DB 143 SVPNFAFGMTFLEINFGAEIYYKHVCHGCGCVFESKISGVLTFMTSPYISIMLC 202
 QY 209 VYRIYLIKEQARLISDANQKLGLEMKNGISQSKERKAVKTLGIWGVFLICWCPFE 268
 DB 203 VYRIYLIKEQARLISDANQKLGLEMKNGISQSKERKAVKTLGIWGVFLICWCPFE 262
 QY 269 ICTVWDPELHYITPPTLNDVLIWFGYLNSTENPMYAFYPMFRKALM 317
 DB 263 ICTVWDPELHYITPPTLNDVLIWFGYLNSTENPMYAFYPMFRKALM 311
 RESULT 14
 ID AAG80972 standard; protein, 296 AA.
 XX
 XX AAG80972;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 XX Human nGPCR56 #2.
 DE
 XX
 KM G protein-coupled receptor; nGPCR; seven transmembrane receptor;
 KM signal transduction; schizophrenia; thyroid disorder; renal failure;
 KM rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KM cardiovascular disease; proliferative disorder; hormonal disorder;
 KM neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KM attention deficit-hyperactivity disorder/attention deficit disorder;
 KM Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KM rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KM neuroprotective.
 KM
 XX
 OS Homo sapiens.
 XX
 XX WO200136473-A2.
 PN
 XX
 PD 25-MAY-2001.
 XX
 PF 16-NOV-2000; 2000WO-US031581.
 XX
 PR 16-NOV-1999; 99US-0165838P.
 PR 17-NOV-1999; 99US-0166071P.
 PR 19-NOV-1999; 99US-0166678P.
 PR 28-DEC-1999; 99US-0173396P.
 PR 22-FEB-2000; 2000US-0184129P.
 PR 28-FEB-2000; 2000US-0186811P.
 PR 28-FEB-2000; 2000US-0185542P.
 PR 02-MAR-2000; 2000US-0186530P.
 PR 03-MAR-2000; 2000US-0186811P.
 PR 09-MAR-2000; 2000US-0188114P.
 PR 17-MAR-2000; 2000US-0190310P.
 PR 21-MAR-2000; 2000US-0190800P.
 PR 20-APR-2000; 2000US-0198568P.
 PR 02-MAY-2000; 2000US-0201190P.
 PR 08-MAY-2000; 2000US-0203111P.
 PR 25-MAY-2000; 2000US-0207094P.
 XX
 XX (PRLA) PHARMACIA & UPJOHN CO.
 PA
 XX

PI Vogel I G, Wood LS, Parodi LA, Hiebach RR, Lind P, Slighcom J;
PI Schellin KA, Kayles PS, Bannigan CW, Ruff V, Sejltitz T, Huff RM;
XX
DR WPI: 2001-389826/41.
XX
DR N-PSDB; AAH51012.
XX
PT New G protein-coupled receptor (ngPCR-x) and its encoding polynucleotide
PT useful for diagnosing and treating e.g. schizophrenia.
XX
PS Claim 37; Page 91-92; 261pp; English.
XX
XX The present invention relates to novel G protein-coupled receptors
CC (ngPCR; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27, 28,
CC 31-38, 40, 41, 53-60) and their coding sequences. The present sequence is
CC one such G protein-coupled receptor. GPCRs are also known as seven
CC transmembrane receptors and function in signal transduction. The ngPCR
CC coding sequences are useful for screening a human to diagnose a disorder
CC affecting the brain or a genetic predisposition, specifically
CC schizophrenia. ngPCR are useful for identifying compounds useful for
CC treating schizophrenia. Detection of ngPCR in a sample is useful as a
CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
CC metabolic and cardiovascular diseases, proliferative disorders and
CC hormonal disorders. Modulators of ngPCR activity have the utility for
CC treating neurological disorders, including schizophrenia, ADHD/ADD
CC (attention deficit-hyperactivity disorder/attention deficit disorder),
CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
CC migraine and senile dementia. Additional disorders include inflammatory
CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
CC diseases e.g. inflammatory bowel disease
XX
SQ Sequence 296 AA;

Query Match 84.4%; Score 1525; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.2e-164;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPFCHNIINISCVKNNMSNDVRASLYSLMWLIITLVGNLIIVVISISHPKQLHTPTNW 60
DB 1 MHPFCHNIINISCVKNNMSNDVRASLYSLMWLIITLVGNLIIVVISISHPKQLHTPTNW 60
QY 61 LIHSNATVDFLGLCLVMPYSWVRSAEHCWYFGEVFCIKHTSTDIMSSASIFHLFSISID 120
DB 61 LIHSNATVDFLGLCLVMPYSWVRSAEHCWYFGEVFCIKHTSTDIMSSASIFHLFSISID 120
QY 121 RYVAVCDDPLRYAKKNNIIVICWIFISWSVPAVFAFGMIFLELNFGAEIYYKVVHCRG 180
DB 121 RYVAVCDDPLRYAKKNNIIVICWIFISWSVPAVFAFGMIFLELNFGAEIYYKVVHCRG 180
QY 181 GGSVFPKISGVLTFMTSFYIPGSIIMLCVYRIRIYLAKEQARLISDANQKLIQIGLEMKNG 240
DB 181 GGSVFPKISGVLTFMTSFYIPGSIIMLCVYRIRIYLAKEQARLISDANQKLIQIGLEMKNG 240
QY 241 ISQSKERKAVKTLGIYMGVFLICWCPFFICTWMDPELHIIPTLND 287
DB 241 ISQSKERKAVKTLGIYMGVFLICWCPFFICTWMDPELHIIPTLND 287
RESULT 15
ABG93790
ID ABG93790 standard; protein; 296 AA.
XX
AC ABG93790;
XX
XX 26-NOV-2002 (first entry)
XX
DE Human G protein-coupled receptor protein, ngPCR-56, #1.
XX
XX Human; receptor; G protein-coupled receptor; GPCR; ngPCR; beGPCR;
KW ng protein coupled receptor; communication; serpentine structure;
KW seven transmembrane receptor; 7TM; mental disorder; diagnosis;
KW genetic predisposition; brain; immune response; gene therapy;

KW anxiety disorder; depression; bipolar disorder; schizophrenia;
KW Huntington's disease; dyskinesia; manic depression; stroke;
KW Parkinson's disease; Alzheimer's disease; diabetes; inflammation; wound;
KW tranquilliser.
XX
OS Homo sapiens.
XX
PN WO200264789-A1.
XX
XX 22-AUG-2002.
XX
XX 14-FEB-2001; 2001WO-US004641.
XX
XX 14-FEB-2001; 2001WO-US004641.
XX
XX (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Lind P, Parodi LA, Vogel I G, Wood LS;
PI
XX WPI: 2002-674879/72.
XX
DR N-PSDB; AB870245.
XX
XX New nucleic acids and polypeptides of the ng protein-coupled receptor,
PT useful for treating or diagnosing a mental disorder or a disorder
PT affecting the brain, e.g. anxiety disorders, schizophrenia, stroke or
PT Parkinson's disease.
XX
PS Example 1; Page 86-87; 244pp; English.
XX
XX The invention discloses an isolated human polypeptide, and encoding
CC nucleic acid, for a G protein-coupled receptor (GPCR), particularly the
CC ng protein coupled receptor-14 (ngPCR-14). GPCRs are vital in the
CC communication between cells and their environment and are characterised
CC by a serpentine structure that passes through the cell membrane seven
CC times, hence the reason such receptors are sometimes called seven
CC transmembrane receptors (7TM). The polynucleotides and polypeptides are
CC useful for identifying an ngPCR allelic variant that correlates with a
CC mental disorder, for isolating an antibody that binds to an epitope of
CC the polypeptide, for identifying a compound that binds the polypeptide or
CC polynucleotide and/or modulates its biological activity, for screening a
CC human subject to diagnose a disorder, or a genetic predisposition to a
CC disorder, affecting the brain or a genetic disposition to the disorder,
CC for identifying compounds useful for the treatment of a mental disorder
CC and for identifying a compound useful as a modulator of binding between
CC ngPCR-14 and a binding partner of ngPCR-14. The polypeptide is also
CC useful for inducing an immune response in a mammal. The nucleic acid or
CC polypeptide is particularly useful, using gene therapy, for treating e.g.
CC anxiety disorders, depression, bipolar disorder, schizophrenia,
CC Huntington's disease, dyskinesias, manic depression, stroke, Parkinson's
CC disease or Alzheimer's disease. The nucleic acid and polypeptide may also
CC be used for treating diabetes, inflammation or wounds. The sequences
CC presented in ABG93747-ABG93793, ABG93795 and ABG93796 are the ngPCR (also
CC referred to as beGPCR) proteins
XX
SQ Sequence 296 AA;

Query Match 84.4%; Score 1525; DB 5; Length 296;
Best Local Similarity 100.0%; Pred. No. 1.2e-164;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPFCHNIINISCVKNNMSNDVRASLYSLMWLIITLVGNLIIVVISISHPKQLHTPTNW 60
DB 1 MHPFCHNIINISCVKNNMSNDVRASLYSLMWLIITLVGNLIIVVISISHPKQLHTPTNW 60
QY 61 LIHSNATVDFLGLCLVMPYSWVRSAEHCWYFGEVFCIKHTSTDIMSSASIFHLFSISID 120
DB 61 LIHSNATVDFLGLCLVMPYSWVRSAEHCWYFGEVFCIKHTSTDIMSSASIFHLFSISID 120
QY 121 RYVAVCDDPLRYAKKNNIIVICWIFISWSVPAVFAFGMIFLELNFGAEIYYKVVHCRG 180
DB 121 RYVAVCDDPLRYAKKNNIIVICWIFISWSVPAVFAFGMIFLELNFGAEIYYKVVHCRG 180
QY 181 GGSVFPKISGVLTFMTSFYIPGSIIMLCVYRIRIYLAKEQARLISDANQKLIQIGLEMKNG 240
DB 181 GGSVFPKISGVLTFMTSFYIPGSIIMLCVYRIRIYLAKEQARLISDANQKLIQIGLEMKNG 240

Db	181	GCSVFPSKISGVLFMTSFYIPGSIMLCVYYRIYLAKEQARLISDANOQLQIGLEMKNG	240
Qy	241	ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLND	287
Db	241	ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLND	287

Search completed: February 12, 2005, 03:43:20
Job time : 78 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 03:33:53 ; Search time 24 Seconds
(without alignments)
1054.418 Million cell updates/sec

Title: US-09-980-145-6
Perfect score: 1806
Sequence: 1 MMPCHNIIINISCVKNMNSN.....FGKIPQDSRCKLFLFLSS 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6C_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1791	99.2	339	4	US-09-659-519-2
2	1393	77.1	332	4	US-09-659-519-4
3	813.5	45.0	343	2	US-08-788-539A-2
4	782	43.3	338	3	US-09-286-805-4
5	743	41.1	348	3	US-09-286-805-2
6	677.5	37.5	337	2	US-08-467-559B-2
7	534	29.6	360	4	US-09-555-313B-4
8	534	29.6	380	4	US-09-555-313B-24
9	534	29.6	388	4	US-08-446-822-8
10	534	29.6	388	3	US-09-328-314-8
11	534	29.6	388	5	PCT-US93-12586-8
12	534	29.6	760	4	US-09-555-313B-2
13	532	29.5	360	4	US-09-826-509-443
14	532	29.5	378	4	US-09-826-509-445
15	532	29.5	380	4	US-09-826-509-441
16	532	29.5	387	4	US-09-826-509-436
17	532	29.5	388	4	US-09-826-509-439
18	529	29.3	387	1	US-07-996-772A-2
19	529	29.3	387	1	US-08-446-822-2
20	529	29.3	387	1	US-09-328-314-2
21	529	29.3	387	5	PCT-US93-12586-2
22	529	29.3	406	5	US-09-328-314-4
23	527	29.2	406	1	US-07-996-772A-4
24	527	29.2	406	1	US-08-446-822-4
25	527	29.2	406	5	PCT-US93-12586-4
26	486	26.9	358	1	US-08-748-485-6
27	478.5	26.5	376	1	US-07-817-920-5

28	478.5	26.5	376	1	US-08-117-006-5	Sequence 5, App11
29	478.5	26.5	376	1	US-08-216-594-5	Sequence 5, App11
30	478.5	26.5	376	5	PCT-US93-00149-5	Sequence 5, App11
31	477.5	26.4	375	1	US-08-370-542-5	Sequence 5, App11
32	477.5	26.4	375	1	US-08-542-358-5	Sequence 5, App11
33	477.5	26.4	375	3	US-09-018-351-5	Sequence 5, App11
34	474.5	26.3	359	3	US-08-875-540-15	Sequence 15, App1
35	474.5	26.3	359	3	US-09-473-634-15	Sequence 15, App1
36	474	26.2	377	2	US-08-461-812-2	Sequence 2, App11
37	474	26.2	377	2	US-08-157-185-14	Sequence 14, App1
38	474	26.2	377	3	US-08-281-526B-14	Sequence 14, App1
39	474	26.2	377	4	US-09-450-790A-14	Sequence 14, App1
40	474	26.2	377	4	US-09-332-837-14	Sequence 14, App1
41	474	26.2	377	4	US-09-371-705-2	Sequence 2, App11
42	472.5	26.2	359	2	US-08-467-568-13	Sequence 13, App1
43	472.5	26.2	359	2	US-08-748-485-4	Sequence 4, App11
44	472.5	26.2	359	2	US-08-103-170-2	Sequence 2, App11
45	472.5	26.2	359	2	US-09-030-582-13	Sequence 13, App1

ALIGNMENTS

```
RESULT 1
US-09-659-519-2
; Sequence 2, Application US/09659519
; Patent No. 6783973
; GENERAL INFORMATION:
; APPLICANT: Grandy, David K
; APPLICANT: Bunzow, James R
; TITLE OF INVENTION: Mammalian Catecholamine Receptor Genes and Uses
; FILE REFERENCE: Catecholamine receptor genes
; CURRENT APPLICATION NUMBER: US/09/659,519
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-659-519-2
```

Query Match 99.2%; Score 1791; DB 4; Length 339;
Best Local Similarity 99.4%; Pred. No. 1.3e+149;
Matches 337; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	MMPCHNIIINISCVKNMNSNDYRASLYSLMTLITLTVGNLIIVYSISHFQQLHTPTNW	60
DB	1	MMPCHNIIINISCVKNMNSNDYRASLYSLMTLITLTVGNLIIVYSISHFQQLHTPTNW	60
QY	61	LHSMATVDLIGCLVMPYSWRSABHCWYRGEVCKIHTSDIMLSASIFHLSPISID	120
DB	61	LHSMATVDLIGCLVMPYSWRSABHCWYRGEVCKIHTSDIMLSASIFHLSPISID	120
QY	121	RYAAVCDPLRYAKXNIIIVICMIFISWSVPAVFAFGMIFLEINFGABEIIYKHYHCRG	180
DB	121	RYAAVCDPLRYAKXNIIIVICMIFISWSVPAVFAFGMIFLEINFGABEIIYKHYHCRG	180
QY	181	GCSVPSKISGYLTMTSTFYIPGSIIMLCVYRIIYIAKEQARLISDANKOIGLEMKNG	240
DB	181	GCSVPSKISGYLTMTSTFYIPGSIIMLCVYRIIYIAKEQARLISDANKOIGLEMKNG	240
QY	241	ISQSEKRAVKTLGIVMGVFLICWCPFLICTMDDPFLHYIIPPTLNDVLIWGYLNSTFN	300
DB	241	ISQSEKRAVKTLGIVMGVFLICWCPFLICTMDDPFLHYIIPPTLNDVLIWGYLNSTFN	300
QY	301	PMVYAFYFVWFRKALKMLFGKIPOKDSRCKLFLFLSS	339
DB	301	PMVYAFYFVWFRKALKMLFGKIPOKDSRCKLFLFLSS	339
RESULT 2		US-09-659-519-4	

```
; Sequence 4, Application US/09659519
; Patent No. 6783973
; GENERAL INFORMATION:
; APPLICANT: Grandy, David K
; APPLICANT: Bunzow, James R
; TITLE OF INVENTION: Mammalian Catecholamine Receptor Genes and Uses
; FILE REFERENCE: Catecholamine receptor genes
; CURRENT APPLICATION NUMBER: US/09/659,519
; CURRENT FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-659-519-4
```

```
Query Match          77.1%; Score 1393; DB 4; Length 332;
Best Local Similarity 77.5%; Pred. No. 1.2e-114;
Matches 259; Conservative 27; Mismatches 46; Indels 2; Gaps 1;
```

```
QY 2 MPECHNINISCVKNNMSNDVRASLSLWVLIITLVGNLIVYISIFKQIHTPTNWL 61
DB 1 MHLCHSANISHTNRMSRDRASLSLSLITLVGNLIVYISIFKQIHTPTNWL 60
QY 62 IHSMTVDPLGLVMPYSWVRSAEHCWYFGEVFCIKHTSTDIMLSASIFHLSFISDR 121
DB 61 LHSMAVVDPLGLVMPYSWVRSAEHCWYFGEVFCIKHTSTDIMLSASIFHLSFISDR 120
QY 122 YVAVCDPLKXKAKMLIVICWIFISWSPVAFPMITLINFKAETIYKHYHCRGG 181
DB 121 YVAVCDPLKXKAKMLIVICWIFISWSPVAFPMITLINFKAETIYKHYHCRGG 180
QY 182 CSVFESKISGVLTFTMTSPYIPGSIIMLCVYRYRIYLAKEOARLISDANQKLGLEMKGI 241
DB 181 CEFLEFSKVGSLAFMTSPYIPGSIIMLCVYRYRIYLAKEOARLISDANQKLGLEMKGI 238
QY 242 SOSKRAVKITIGIYWGVLICWCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTNPN 301
DB 239 POSKRTKAKITIGIYWGVLICWCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTNPN 298
QY 302 WYVAFYFWFRKALKMLFGKIFQKDSRCKLFL 335
DB 299 WYVAFYFWFRKALKMLFGKIFQKDSRCKLFL 332
```

RESULT 3

```
US-08-788-539A-2
; Sequence 2, Application US/08788539A
; Patent No. 5871967
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: CLONING OF A NOVEL G-PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/788,539A
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50047
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-788-539A-2
```

```
Query Match          45.0%; Score 813.5; DB 2; Length 343;
Best Local Similarity 46.0%; Pred. No. 1e-63;
Matches 159; Conservative 57; Mismatches 99; Indels 31; Gaps 5;
```

```
QY 5 CHNIIIN-----ISCVKNN-----MSNDVRASLSLWVLIITLVGNLIVYISISH 50
DB 13 CPKFNKILSHQPLFSCPDGNVFGYDKSHDY-----PLFGNLIVYISISH 58
QY 51 FKQIHTPTNWL IHSMTVDPLGLVMPYSWVRSAEHCWYFGEVFCIKHTSTDIMLSAS 110
DB 59 FKQIHTPTNWL IHSMTVDPLGLVMPYSWVRSAEHCWYFGEVFCIKHTSTDIMLSAS 118
QY 111 IFHLSFISDRYVAVCDPLKXKAKMLIVICWIFISWSPVAFPMITLINFKAETIYKHYH 170
DB 119 IFHLSFISDRYVAVCDPLKXKAKMLIVICWIFISWSPVAFPMITLINFKAETIYKHYH 178
QY 171 IYVHYHCRGGCVSFESKISGVLTFTMTSPYIPGSIIMLCVYRYRIYLAKEOARLISDANQ 230
DB 179 -YKILVACFNFCALTFNFKWGTILFTTCFPLPGSIIMVIGIKFIYVSKQARVISHPEN 237
QY 231 LQIGLEMKNGISOSKRAVKITIGIYWGVLICWCPFFICTVMDPLHYIIPPTLNDVLI 290
DB 238 TKGA--YKHLISKQDKRAKKTIGIYWGVLICWCPFFICTVMDPLHYIIPPTLNDVLI 295
QY 291 WFGYLNSTNPNWYVAFYFWFRKALKMLFGKIFQKDSRCKLFL 336
DB 296 WLRYSNSTCNPLIHGFENFWQAKFKYIVSGKIFSSHSETANLPPE 341
```

RESULT 4

```
US-09-286-805-4
; Sequence 4, Application US/09286805
; Patent No. 6117990
; GENERAL INFORMATION:
; APPLICANT: Bontini, James A.
; APPLICANT: Borowsky, Beth E.
; TITLE OF INVENTION: DNA Encoding Orphan SNORF1 Receptor
; FILE REFERENCE: 58987
; CURRENT APPLICATION NUMBER: US/09/286,805
; CURRENT FILING DATE: 1999-04-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 4
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-286-805-4
```

```
Query Match          43.3%; Score 782; DB 3; Length 338;
Best Local Similarity 45.8%; Pred. No. 5.9e-61;
Matches 154; Conservative 63; Mismatches 115; Indels 4; Gaps 4;
```

```
QY 2 MPECHNINISCVKNNMSNDVRASLSLWVLIITLVGNLIVYISIFKQIHTPTNWL 61
DB 1 MELCYENVNGSCIKSSYSPWPRALIVAVGIGALLAVFGNLLVITAIHFQIHTPTNWL 60
```

QY 122 YVAVCDPLRYAKMNLIVICMI FISMVPAVAFAGMI FLELNFGABEIIYYKHVCRGG 181

Db 121 YVAVTBLPTPTKFTISVSGVCIALSMFSPVSYTSISITYGTANBEGIEBLVVA -LTCVGG 179

QY 182 CSAFPGKISGVLFEMTISFYIPGSIMLCVYRYLYLAKRQALIL -SDNNOKLQIGLEMKNG 240

Db 180 CQAPLNNWVLLCFEL -FELPTVYVWVFLKGRIFLPAKQAKIRBSAQQPOASESEYKER 238

QY 241 ISGSKERKAVKTLGIWGVFLICWCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTFN 300

Db 239 VAR-RERKAKTIGIAMAFLVSMPLPIIIDAVIDAYNMNFIPTPAYVEILVWCYYNSAMN 297

QY 301 PMYAFPYPMERKALQMMLFGKIPQKDSRCLTLE 336

Db 298 PLIYAFPYPMERKALKLIISGKFRDSSRTINLESE 333

```

RESULT 5
US-09-286-805-2
: Sequence 2 Application US/09286805
: Patent No. 6117990
: GENERAL INFORMATION:
: APPLICANT: Bonini, James A.
: APPLICANT: Borovsky, Beth E.
: TITLE OF INVENTION: DNA Encoding Orphan Smorfl Receptor
: FILE REFERENCE: 58967
: CURRENT APPLICATION NUMBER: US/09/286,805
: CURRENT FILING DATE: 1999-04-06
: NUMBER OF SEQ. ID NOS.: 4
: SOFTWARE: PatentIn Ver. 2.0 - beta
: SEQ. ID NO. 2
: LENGTH: 348
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-286-805-2

```

Query Match	41.1%	Score 743	DB 3	Length 348
Best Local Similarity	42.5%	Pred. No. 1.6e-57		
Matches 141	Conservative	63	Mismatches 126	Indels 2
Gap				2
QY	5	CHNIINISCVKNMNSNDVRASLYSLMVLIIITLTGVGNLIVISISHFKOLHTPTNMLIHS	64	
Db	14	CYKVNNECITKTPYSPGPRSLIYALVAGGVALAARGNLLMALIHFQKHPTNPLIAS	73	
QY	65	MATVDFLLGLVMPYSMWRSAEHCMYFGGEVFCCKHTSTDIMLSASIFHLSTSIDRYA	124	
Db	74	LACADFVLGVATMPSPSTVRSVESCMYFGDSYCKFHFTCEPTSCFPASLFLHCCISIDRYA	133	
QY	125	VCDDPLRYAKNNIIIVICWIFISMSVPAVFAECGMFLDELNFGAREIYKHNHCRCGCGSV	184	
Db	134	VTDFLTPTTKFTVSVSGICIVISMFPSTVYSISITTYGANNGBIELVVA-LTCVGGCOA	192	
QY	185	FFSKISGVLTTMTSTFYIFGSIIMLCVYRIYLIYLAKEQARLISDANOKLQIGLEMKRGISQS	244	
Db	193	PLNQWVWLFCFL-FFIRNVAMVFIYSKIFLVAKKQARKIESTSAQAQSSSYKERVAK	251	
QY	245	KERKAVKTLGIVMGVFLICWCBFFICTYVMDPEFLHYIIPETLNDVLIWFGYLNSTNPMVY	304	
Db	252	RERKAKTLGIAMAFAFLVSMLEPYLDAVIDAYMNETTPPYVEIILVWCYVNSANPLIY	311	
QY	305	AFPYWPRKALKMLFGKIFQDSSRCULFLE	336	
Db	312	AFPYQWFGKAIKLIVSGKVLRTDSSSTYNLFSE	343	

RESULT 6
US-08-467-559B-2
: Sequence 2, Application US/08467559B

```

1 Patent No. 5928890
2
3 GENERAL INFORMATION:
4
5 APPLICANT: LI, YI
6
7 TITLE OF INVENTION: HUMAN AMINE RECEPTOR
8
9 NUMBER OF SEQUENCES: 10
10
11 CORRESPONDENCE ADDRESS:
12
13 ADDRESSEE: STERN, KESSLER, GOLDSTEIN AND FOX, P.L.L.C.
14
15 STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
16
17 CITY: WASHINGTON
18
19 STATE: DC
20
21 COUNTRY: UNITED STATES OF AMERICA
22
23 ZIP: 20005-3934
24
25 COMPUTER READABLE FORM:
26
27 MEDIUM TYPE: Floppy disk
28
29 COMPUTER: IBM PC compatible
30
31 OPERATING SYSTEM: PC-DOS/MS-DOS
32
33 SOFTWARE: PatentIn Release #1.0, Version #1.30
34
35 CURRENT APPLICATION DATA:
36
37 APPLICATION NUMBER: US/08/467,559B
38
39 FILING DATE: 06-JUN-1995
40
41 CLASSIFICATION: 435
42
43 ATTORNEY/AGENT INFORMATION:
44
45 NAME: STEFFE, ERIC K
46
47 REGISTRATION NUMBER: 36,688
48
49 REFERENCE/DOCKET NUMBER: 1488.0840000
50
51 TELECOMMUNICATION INFORMATION:
52
53 TELEPHONE: (202) 371-2500
54
55 TELEFAX: (202) 371-2540
56
57 INFORMATION FOR SEQ ID NO: 2:
58
59 SEQUENCE CHARACTERISTICS:
60
61 LENGTH: 337 amino acids
62
63 TYPE: amino acid
64
65 TOPOLOGY: linear
66
67 MOLECULE TYPE: protein
68
69
70 US-08-467-559B-2

```

[illegible]

```

RESULT 7
US-09-555-313B-4
; Sequence 4, Application US/09555313B
; Patent No. 6506580
; GENERAL INFORMATION:
; APPLICANT: FISCHMEISTER, Rudolph et al.
; TITLE OF INVENTION: Splicing variants of the human serotoninergic receptor
; TITLE OF INVENTION: 5-HT4 and uses thereof, in particular for screening

```

```
FILE REFERENCE: P06762US00/BAS
CURRENT APPLICATION NUMBER: US/09/555,313B
CURRENT FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: FR 97/15037
PRIOR FILING DATE: 1997-11-28
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
US-09-555-313B-4
```

```
Query Match      29.6%; Score 534; DB 4; Length 360;
Best Local Similarity 37.4%; Pred. No. 3.9e-39;
Matches 116; Conservative 53; Mismatches 121; Indels 20; Gaps 5;
```

```
QY 26 LYSIMVLITLTIVGNLIVISISHPKQLH-TPTNMLHSMATVDFLGLCLVMPYSWRS 84
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 22 LITFLSTVILMALILGNLIVMAVAVCMWRQLRKIKTNFYISLAFADLLVSLVWPFGAIEL 81

QY 85 AEHCWYFGEVFCCKHTSTDIMLSASIFHLFSISIDRYAV-CDPLRYKAKNNILVICWM 143
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 82 VQDIWIYGEVFCIVRSIDVLLTTASIFHLCCISLDRYTAICQPLVYRNKMTPLRIALM 141

QY 144 IFTISNVPVAFNAGMIFLELNFGAEIYYKHHVCRG-----CSVFSKISGVLTFMTSF 199
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 LGGCWVIFPTFISFLPMOGMNNIGIIDLEKRFNONSNTYCVPMVKNPYAITSVAVF 201

QY 200 YIPGSIIMLCVYRYIYLAKEQARLISDANOKIOLGLEMKNGISQS-----KERKA 249
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 202 YIPFLMLVATYRIYVTAKEHAHQI---QMLQAGASSESRPOSADQSTHRMRTETKA 257

QY 250 VKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFPMVYAFEPY 309
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 258 AKTLCTIIMGFCCLCMAPFFVTNIVDPFIDYVPGQWTAFLMLGYNGLNPLFYALFLNK 317

QY 310 WFRKRLKML 319
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 318 SFRRAFLIIL 327
```

```
RESULT 8
US-09-555-313B-24
Sequence 24, Application US/09555313B
Patent No. 6506580
GENERAL INFORMATION:
APPLICANT: FICSMIEISTER, Rudolph et al.
TITLE OF INVENTION: Splicing variants of the human neurotrophic receptor
FILE REFERENCE: P06762US00/BAS
CURRENT APPLICATION NUMBER: US/09/555,313B
CURRENT FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: FR 97/15037
PRIOR FILING DATE: 1997-11-28
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 380
TYPE: PRT
ORGANISM: Homo sapiens
US-09-555-313B-24
```

```
Query Match      29.6%; Score 534; DB 4; Length 380;
Best Local Similarity 37.4%; Pred. No. 4.2e-39;
Matches 116; Conservative 53; Mismatches 121; Indels 20; Gaps 5;
```

```
QY 26 LYSIMVLITLTIVGNLIVISISHPKQLH-TPTNMLHSMATVDFLGLCLVMPYSWRS 84
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 22 LITFLSTVILMALILGNLIVMAVAVCMWRQLRKIKTNFYISLAFADLLVSLVWPFGAIEL 81

QY 85 AEHCWYFGEVFCCKHTSTDIMLSASIFHLFSISIDRYAV-CDPLRYKAKNNILVICWM 143
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 LGGCWVIFPTFISFLPMOGMNNIGIIDLEKRFNONSNTYCVPMVKNPYAITSVAVF 201
```

```
DB 82 VQDIWIYGEVFCIVRSIDVLLTTASIFHLCCISLDRYTAICQPLVYRNKMTPLRIALM 141

QY 144 IFTISNVPVAFNAGMIFLELNFGAEIYYKHHVCRG-----CSVFSKISGVLTFMTSF 199
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 LGGCWVIFPTFISFLPMOGMNNIGIIDLEKRFNONSNTYCVPMVKNPYAITSVAVF 201

QY 200 YIPGSIIMLCVYRYIYLAKEQARLISDANOKIOLGLEMKNGISQS-----KERKA 249
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 202 YIPFLMLVATYRIYVTAKEHAHQI---QMLQAGASSESRPOSADQSTHRMRTETKA 257

QY 250 VKTLGIWGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWFGYLNSTFPMVYAFEPY 309
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 258 AKTLCTIIMGFCCLCMAPFFVTNIVDPFIDYVPGQWTAFLMLGYNGLNPLFYALFLNK 317

QY 310 WFRKRLKML 319
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 318 SFRRAFLIIL 327
```

```
RESULT 9
US-08-446-822-8
Sequence 8, Application US/08446822
Patent No. 5766879
GENERAL INFORMATION:
APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: DNA ENCODING 5-HT4 SEROTONIN RECEPTORS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: COOPER & DUNHAM
STREET: 30 ROCKEFELLER PLAZA
CITY: NEW YORK
STATE: NEW YORK
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,822
FILING DATE: June 1, 1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: White, P., John
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 42667-A-PCT-US/DPW/MAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-822-8
```

```
Query Match      29.6%; Score 534; DB 1; Length 388;
Best Local Similarity 37.4%; Pred. No. 4.3e-39;
Matches 116; Conservative 53; Mismatches 121; Indels 20; Gaps 5;
```

```
QY 26 LYSIMVLITLTIVGNLIVISISHPKQLH-TPTNMLHSMATVDFLGLCLVMPYSWRS 84
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 22 LITFLSTVILMALILGNLIVMAVAVCMWRQLRKIKTNFYISLAFADLLVSLVWPFGAIEL 81

QY 85 AEHCWYFGEVFCCKHTSTDIMLSASIFHLFSISIDRYAV-CDPLRYKAKNNILVICWM 143
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 82 VQDIWIYGEVFCIVRSIDVLLTTASIFHLCCISLDRYTAICQPLVYRNKMTPLRIALM 141

QY 144 IFTISNVPVAFNAGMIFLELNFGAEIYYKHHVCRG-----CSVFSKISGVLTFMTSF 199
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 LGGCWVIFPTFISFLPMOGMNNIGIIDLEKRFNONSNTYCVPMVKNPYAITSVAVF 201
```



```

; FILE REFERENCE: P06762US00/BAs
; CURRENT APPLICATION NUMBER: US/09/555,313B
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: FR 97/15037
; PRIOR FILING DATE: 1997-11-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-555-313B-2

```

Query Match	29.6%;	Score 534;	DB 4;	Length 760;
Best Local Similarity	37.4%;	Pred. No. 8.9e-39;		
Matches 116;	Conservative	53;	Mismatches 121;	Indels 20;
			Gaps	5

```

Oy      26 LYSMLVLLITLTGNLIIVSISHPQLH-TPTNMLIHSNATVPLGLGLMPYSWRS 84
         | : : | : | : | : | : | : | : | : | : | : | : | :
Db      22 LITFLSTVIMAIIGNLWVAVCWDRQLKIKITNYFIYSLAFADLLVSVALMPFGAIEL 81

```

```

QY      85 AEHCWYFGSEVCKLHTSTDMLTSSASIFHLFSISIDRYAV-CDPLRYKAGNNILIVCM 143
      : | : | | | : | | : : : | | | | | | | | : | | | : |
Db      82 VQDIIWYGEVFCVLRTSLDVLTLTASIFHLCCISIDRYAIVCCQPLVYRNKWTPLRIALM 141

```

Dy
144 LFTSMVAVAFARGLFLELNFGAEIYKHVRGG-----CSYFFSKISGVLTFMTSF 199

Dd
142 LGSCWLPFLPISFLPMGMNNTIGIDLEKRFENQNSNSTYCVPMMNKPEYAITSVAF 201

Db 202 YIPFLIMVLAAYRIYVTAKEHAHQI---QMLQAGAGASSSRPGADQHSHTRMKTETKA 257

D5
258 AKTLCIMGCFCTCWAPFVNTINVDPIFYTVPGQVWTAFLMLGINSGLNPLDYAFLNK 317

Db 318 SERRAFALLI 327

RESULT 13
US-09-826-509-443
; Sequence 443, Application US/09826509
Parent No. 00000000

```

; GENERAL INFORMATION:
;
; APPLICANT: Lehmann-Brulnsma, Karin
;
; APPLICANT: Liaw, Chen W.
;
; APPLICANT: Lin, Tien

```

TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
 TITLE OF INVENTION: Protein-Coupled Receptors
 FILE REFERENCE: AREN-207
 CURRENT APPLICATION NUMBER: US/09/0826 500

/ CONVENT FILING DATE: 2001-04-05
 / PRIOR APPLICATION NUMBER: 60/195,747
 / PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: 09/170,496

```

; NUMBER OF SEQ ID NOS: 589
;
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 443

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-443

```

Query Match	29.5%;	Score 532;	DB 4;	Length 360;
Best Local Similarity	37.4%;	Pred. No. 5.9e-39;		
Matches 116;	Conservative	53;	Mismatches 121;	Indels 20;
			Gaps	5;

26 LYSMLVLLITLVGNLIVIVISISHKQLH-TPTNWLHSMATVDFLLGLVMPYSWRS 84

```

Db      22  LTFEJSTVJIMALIGRLVMAVCMWDQRLKTNPIFVSLAFADLLVSLVMPFGAIEL  81
QY      85  AHCWCFGEFPCXIHSTIDMLSSASTHSLFISIDRYAV-COPLRYKAKOMLIV  CVM 143
Db      82  VOIMTMYGEFPCILVRISDLVLTTSASIFHLCISIDRYAICCPPLRYRKMPLRLALM 141
QY      144  IFTSWSVPAVFAFGMIFLELNFKAAGEIYYGHVHCROG---CSVPEFSKISGVLTFWTSF 199
Db      142  LGGCWVYIPPIISPLPIMGOMNIGIIDLERKXNQNQNSNTRYCYPMWNNKYAITCSYAF 201
QY      200  YIPGSIMLCVYRIYILIAEQARLISDANOKLQIGLEMKGISOS-----KERRA 249
Db      202  YIFPLMLVAYRIYVLTAEHQAQI---QMLQAGASSBSRPSADQHSHTHRMRTETYA 257
QY      250  VKTLGIWGVPLICCPFFICTWMDPELHYIIPFLINDVILIFGYNLSTENPMUYAFYF 309
Db      258  KKTLCIIMGCFGLCANAPFVTNIYDPIFYVPGQWTAFLMLGIYNSGLNAPPLXAFNFK 317
QY      310  WPKKALXOML 319
Db      318  SFRRAFLITL 327

```

RESULT 14
US-09-826-509-445

```

; Patent No. 6806054
;
; GENERAL INFORMATION:
;
; APPLICANT: Lehmann-Bruhnsma, Karin
;
; APPLICANT: Frau Chob W

```

```

: APPLICANT: Lin, I-Lin
: TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known C
: TITLE OF INVENTION: Protein-Coupled Receptors
: FILE REFERENCE: APMN-2007

```

:
 : CURRENT APPLICATION NUMBER: US/09/826,509
 :
 : CURRENT FILING DATE: 2001-04-05
 :
 : PRIOR APPLICATION NUMBER: 60/195,747
 :
 : PRIOR FILING DATE: 2000-04-07
 :

? REASON FOR REJECTION NUMBER: 05/11/01, 450
 ? PRIOR FILING DATE: 1998-10-13
 ? NUMBER OF SEQ ID NOS: 589
 ? SOFTWARE: PatentIn Version 2.1

```

; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens

```

Query Match	29.5%;	Score 532;	DB 4;	Length 378;
Best Local Similarity	37.4%;	Pred. No. 6.2e-39;		

26 LYSIMWLIIITLVGNLIVIVISISHRQOLH-TPTNWLHSMATVDFLLGCLVMPYSMYS 84

[illegible]

QY 144 IFISMSVPAVEAFGMFLEINPKGAEIYYHVHCRGG---CSVFESKISGVLTFTMTSF 199
:
: : : : :
Nb 142 IGGCWITPTETSELPFMCGANNIGTIDLTERPRENONSSTVCUMENNRDVAITCCGTTATF 201

Qy 200 YIPGIMLCVYYRYIYIAKEQARLISDANOQIÖIGLEMKNGISÖS-----KERKA 249
||| :|: ||||: ||| |
Db 202 YIPFLIMVLAYYRIYTTAKKAHOT-----OMTÖPAGASSFSBPDSADOSHTEHPMTETKA 257
||| :|: ||||: ||| |

QY
250 VKTLGIWGVFLICWCPEFICTVMDPRLHYITIPPLINDVLTWPGYNLSTENPMVYAFFEP 309
|||::||:||||:||||:|:||:|||||
258 KTLTCTIMGEFCICWAPEFVNIVPEIDYWPQGWATAPLMGYNSGIDNPETYYAFINR 317
|||::||:||||:||||:|:||:|||||

QY 310 WFRKALKML 319

```

Db      22  LTFEJSTVJIMALIGRLVMAVCMWDQRLKINPIFVSLAFADLLVSVMPFGAIEL  81
QY      85  AHCWCFGEFPCXIHSTIDMLSSASTHSLFISIDRYAV-COPLRYKAKOMLIV  CVM 143
Db      82  VOIMTMYGEFPCXIVRSIDLVLLTTTASIFHLCCISIDRYAACCPPLRYRKMPLRLALM 141
QY      144  IFTSWSVPAVFAFGMIFLELNFKAAGEIYYGHVHCROG---CSVPEFSKISGVLTFWTSF 199
Db      142  LGGCWVIPIPIISPLPIMGOMNIGIIDLERKXNQNQNSNTRYCYPMWNNKYAITCSYAF 201
QY      200  YIPGSIMLCVYRIYILIAEQARLISDANOKLOGLEMKGISOS-----KERRA 249
Db      202  YIFPLMLVAYRIYVLTAEHQAQI---QMLQAGASSBSRPSADQHSHTHRMRTETYA 257
QY      250  VKTLGIWGVPLICCPFFICTWMDPELHYIIPFLINDVILIFGYNLSTENPMUYAFYF 309
Db      258  KKTLCIIMGCFCICNAPFPVTNIYDPIFYVPGQWTAFLMLGIYNSGLNAPPLXAFNFK 317
QY      310  WPKKALXOML 319
Db      318  SFRRAFLITL 327

```

RESULT 14
US-09-826-509-445

```

; Patent No. 6806054
;
; GENERAL INFORMATION:
;
; APPLICANT: Lehmann-Bruhnsma, Karin
;
; APPLICANT: Frau Chob W

```

```

: APPLICANT: Lin, I-Lin
: TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known C
: TITLE OF INVENTION: Protein-Coupled Receptors
: FILE REFERENCE: APMN-2007

```

:
 : CURRENT APPLICATION NUMBER: US/09/826,509
 :
 : CURRENT FILING DATE: 2001-04-05
 :
 : PRIOR APPLICATION NUMBER: 60/195,747
 :
 : PRIOR FILING DATE: 2000-04-07
 :

? REASON FOR REJECTION NUMBER: 05/11/01, 450
 ? PRIOR FILING DATE: 1998-10-13
 ? NUMBER OF SEQ ID NOS: 589
 ? SOFTWARE: PatentIn Version 2.1

```

; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens

```

Query Match	29.5%;	Score 532;	DB 4;	Length 378;
Best Local Similarity	37.4%;	Pred. No. 6.2e-39;		

26 LYSIMWLIIITLVGNLIVIVISHSRQOLH-TPTNWLHSMATVDFLLGCLVMPYSMYS 84

```

QY      85 AEHCWFEVCHKIHSTDIMLSASIFHLSPISIDRYAV-CDPLRYAKNNILVICM 143
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
b      82 VQDILVVCSEVCTIDMETDUTITMELQFURCOTDPRVATGQOYDREWDVETLVYV 144

```

QY 144 IFISMSVPAVEAFGMFLEINPKGAEIYYHVHCRGG---CSVFESKISGVLTFTMTSF 199
:
: : : : :
Nb 142 IGGCWITPTETSELPFMCGANNIGTIDLTERPRENONSSTVCUMENNRDVAITCCGTTATF 201

```
QY 200 YIPGIMLCVYYRYIYIAKEQARLISDANOXIOGLEMKNGISQS-----KERKA 249
||| :|: ||||: ||| |
Db 202 YIPFLIMVLAYYRIYTTAKKAHOT-----ONT-OPAGASSFSBDOASDHSTHMPTEKKA 257
||| :|: ||||: ||| |
```

QY
250 VKTLGIWGVFLICWCPEFICTVMDPRLHYITIPPLINDVLTWPGYNLSTENPMVYAFFEP 309
|||::||:||||:||||:|:||:|||||
258 KTLTCTIMGEFCICWAPEFVNIVPEIDYPWGOWATAFLMIGYNSGIDNPETYYAFINR 317
|||::||:||||:||||:|:||:|||||

QY 310 WFRKALKMML 319

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 03:30:08 ; Search time 22 Seconds
(without alignments)
1482.613 Million cell updates/sec

Title: US-09-980-145-6
Perfect score: 1806
Sequence: 1 MPMPCNNINISCVKNNMSN.....FGKIFQKDSRCKLFLRLSS 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	677.5	37.5	337	2	JC5832 neurotransmitter r
2	529	29.3	387	2	S55550 5-HT4s receptor -
3	529	29.3	406	2	S55549 serotonin 4 recept
4	510.5	28.3	459	2	A56849 dopamine receptor-
5	497.5	27.5	465	2	I51661 dopamine D1C recep
6	493.5	26.3	363	2	I50475 dopamine D1 recep
7	486	26.9	358	2	J01278 histamine H2 recep
8	482	26.7	444	2	C55886 dopamine receptor
9	482	26.7	450	2	A55886 dopamine receptor
10	481	26.6	457	2	I51660 dopamine D1B recep
11	480	26.6	451	2	I51659 dopamine D1A recep
12	475	26.3	377	2	S68423 serotonin receptor
13	474	26.2	377	2	A53279 serotonin receptor
14	473	26.2	374	2	I77467 serotonin receptor
15	472.5	26.2	359	2	JH0449 histamine H2 recep
16	467	25.9	446	2	I47217 dopamine receptor
17	466.5	25.8	377	2	B30341 G protein-coupled
18	466.5	25.8	386	2	S72168 dopamine receptor
19	462.5	25.6	463	2	B56849 dopamine receptor-
20	461	25.5	483	2	A25896 beta-adrenergic re
21	459.5	25.4	359	2	JC4120 histamine H2 recep
22	458	25.4	446	1	DYH0D1 dopamine receptor
23	454.5	25.2	477	1	DYH0D5 dopamine receptor
24	452	25.0	418	1	ORHYB2 beta-2-adrenergic
25	451	25.0	418	2	S10855 beta-2-adrenergic
26	450.5	24.9	359	2	A39008 histamine H2 recep
27	450	24.9	475	2	A41271 dopamine receptor
28	449.5	24.9	464	2	S12591 beta-1-adrenergic
29	449	24.9	418	2	S00260 beta-2-adrenergic

30	446.5	24.7	415	2	I53040 beta-2 adrenergic
31	446.5	24.7	487	1	DYR7D1 dopamine receptor
32	446	24.7	466	2	S36794 beta-1-adrenergic
33	438.5	24.3	413	1	ORHUB2 beta-2-adrenergic
34	437.5	24.2	486	2	B55886 dopamine receptor
35	433.5	24.0	477	1	QRHUB1 beta-1-adrenergic
36	432.5	23.9	366	2	A47321 serotonin receptor
37	429.5	23.8	428	2	A55044 beta-4C-adrenergic
38	429.5	23.8	477	2	S71323 alpha-1A adrenergic
39	429	23.8	480	2	I53053 beta-1 adrenergic
40	427.5	23.7	366	2	S26048 serotonin receptor
41	427.5	23.7	386	2	A42688 serotonin receptor
42	427.5	23.7	386	2	S18637 serotonin receptor
43	425.5	23.6	429	2	S65656 alpha-1C-adrenergic
44	425.5	23.6	466	2	JN0765 alpha-1C-adrenergic
45	425.5	23.6	499	2	S65657 alpha-1C-adrenergic

ALIGNMENTS

RESULT 1

JC5832
neurotransmitter receptor - human
C/Species: Homo sapiens (man)
C/Date: 20-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C/Accession: JC5832
R/Zeng, Z.; Fan, P.; Rand, E.; Kyaw, H.; Su, K.; Madike, V.; Carter, K.C.; Li, Y.
Biochem. Biophys. Res. Commun. 242, 575-578, 1998
A/Title: Cloning of a putative human neurotransmitter receptor expressed in skeletal mus
A/Reference number: JC5832; MUID:98125534; PMID:9464258
A/Accession: JC5832
A/Molecule type: mRNA
A/Residues: 1-337 <ZEN>
A/Cross-references: UNIPROT:O14604; GB:AF021818; NID:92465431; PIDN:AAC39581.1; PID:9246
A/Experimental source: brain
A/Genetics:
A/Map position: 6q23
C/Keywords: glycoprotein; receptor; transmembrane protein
F/29-63/Domain: transmembrane #status predicted <TM1>
F/69-95/Domain: transmembrane #status predicted <TM2>
F/114-118/Domain: transmembrane #status predicted <TM3>
F/149-173/Domain: transmembrane #status predicted <TM4>
F/186-200/Domain: transmembrane #status predicted <TM5>
F/204-229/Domain: transmembrane #status predicted <TM6>
F/253-310/Domain: transmembrane #status predicted <TM7>
F/21/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.5% Score 677.5; DB 2; Length 337;
Best Local Similarity 39.0% Pred. No. 1.3e-49;
Matches 130; Conservative 67; Mismatches 125; Indels 11; Gaps 4;

QY	4	FCHNINISCVKNNMSNDVRASLYSLMWLIITLVGNLIYIVSISHPKQHTPTNMLIH 63	
DB	16	FCYQ-VNGSCPRVHTTGLIVLYLTCAAGMLIYGVVPAFAVSKALHTPTNFIIL 74	
QY	64	SMATYDFLLGLCLNMPYSNVAENHGWYGEVPCKHSTIDMLSAISFHLSTIDRY 123	
DB	75	SLALDMFLGLVLPSTIRSVESCFPGDLCRIHTYLDLFCUTSIFHLCFSTIDBHC 134	
QY	124	AVCDPLRYKAKNNILVICWMIFFISWVPAVAFGMIFELNPKGABEYVKKHVGSGCS 183	
DB	135	ALCDPLLYPSKFTYRVALRYLLAGNGVPAATSLFLYTDVETRLSQ-WLEBMPGVSCQ 193	
QY	184	VFFSKISGVLTFTMSFYIPGSIIMLCVYRYIYLAKEQARLISDANQKLGLEMGKISQ 243	
DB	194	LILNKFWGMNLF-PLFFVCLIMISLVYKIFVVAATRAQQTTLTSKSP-----AGAA 244	
QY	244	SKERAVYTLGIVNGVFLICWCPFCICVMPPLHYIIPPLNDVLIIFGLINSTFPMV 303	
DB	245	KHERAAATLTGIIVGIVLLCWLPTITDMVDSLHFIIPLVDFIFWFAVFNACNPDI 304	

```

Qy      304 YAFYYPWERKALKMFLGKIQKDSNRCKLFLE 336
      | | | | | | | | | | : : : : |
Db      305 YVFSYQWFRKALKLTLSQKVFSPQTRVLDLYQE 337

```

RESULT 2

5-HT4S receptor - rat
S/Species: Rattus norvegicus (Norway rat)
C/Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C/Accession: S55550
R/Gerald, C., Adham, N., Kao, H.T., Olsen, M.A., Laz, T.M., Schechter, L.E., Bard, J.A.,
EMBO J. 14, 2806-2815, 1995
A/Title: The 5-HT(4) receptor: molecular cloning and pharmacological characterization of
A/Reference number: S55549; MUID:95317299; PMID:7796807
A/Accession: S55550
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-387 <GB>
A/Cross-References: UNIPROT:O62758; GB:U0906; NID:G924638; PIDN:MAC5232.1; PID:G924639
A/Superfamily: vertebrate rhodopsin

Query Match	29.3%;	Score 529;	DB 2;	Length 387;
Best Local Similarity	35.4%;	Pred. No. 4.6e-37;		
Matches 111; Conservative	61;	Mismatches 114;	Indels 28;	Gaps 6

Qy	26	LYSIALMLILITLTVGLVLIYVISHKOLH-TPNMLHSMATYDFLLGLCYMYSVRS	84
Db	22	LITFPAMVILMALIGMLVMAVCHDRÖRKTKTYFVLSVAPDLVSVYNAFGAIEL	81
Qy	85	ABHCWYFGEVCKIHTSTDIMLSASIPHLSPISIDRYAAV-CDPLRYKAMNIVLCVM	143
Db	82	VODIMWYGMFCIVLTSIDLVTLTSTASIPHLCSISDRYYALICQDPLVYRNKMPRLTALM	141
Qy	144	IFISMSVPAVPAFGMIFELNFKGABEY----YCHNCRGCGSVFSPKISGUTIMTST	199
Db	142	LGSCWVIMPFISFLPMQGMNIGIVDYIEKKKFMHNSSTFCVPMWKKPALTCSVAF	201
Qy	200	YIPGSIIMLCVYRIYILAKEQARLISDANKÖKIGIELEMKGISQK-----	245
Db	202	YIPFLMLVLAAYRIIVYTAKEHA-----QOIQM-LQAGATISBSRPOTADQSHTRMRT	253
Qy	246	EKKAVTLIGIVNGVFELICMCPFEICTWMDPFLHYIIPPLNDVLIMNGYLNSTFNPMVYA	305
Db	254	EIKAKATLCVIMGCCFCPCNAPEFVINIVDPFDIYVPEKWTAFELMGIYINSGLNPLVA	313
Qy	306	FFYPWFRKALKMML	319
Db	314	FLNKSEFRRAFLIIL	327

RESULT 3

serotonin 4 receptor - rat
N.Alternate names: 5-HT receptor 4L
C.Species: Rattus norvegicus (Norway rat)
C.Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
R.Accession: S55549; S66494
R.Gerald C., Adnan, N.; Kao, H.T.; Olsen, M.A.; Laz, T.M.; Schechter, L.E.; Bard, J.A. #
EMBO J. 14, 2806-2815, 1995
A.Title: The 5-HT(4) receptor: molecular cloning and pharmacological characterization of
A.Reference number: S55549; MUID:95317299; PMID:7796807
A.Accession: S55549
A.Status: Preliminary
A.Molecule type: mRNA
A.Residues: 1-406 <GER>
R.Cross-references: UNIPROT:O62758; GB:U20907; NID:9924640; PIND:MACS2233.1; PID:9924641
R.Ullmer, C.; Schmuck, K.; Kalkman, H.O.; Luebbert, H.
FEBS Lett. 370, 215-221, 1995
A.Title: Expression of serotonin receptor mRNAs in blood vessels.
A.Reference number: S66487; MUID:95385798; PMID:7656980
A.Accession: S66494
A.Status: nucleic acid sequence not shown

A: Molecule type: mRNA
A: Residues: 95-259 <URL>
A: Cross-references: EMBL:Z48153; NID:G984171; PID:G984172Z1
A: Experimental source: tissue brain
C: Superfamily: vertebrate rhodopsin
C: Keywords: neurotransmitter receptor

Query Match	29.3%;	Score 529,	DB 2;	Length 406;
Best Local Similarity	35.4%;	Pred. No. 4.8e-37;		
Matches 111;	Conservative 61;	Mismatches 114;	Indels 28;	Gaps 6

Qy	26	LYSLAWLIITTLVNGVLIVYISLHFKOLH-TPNMLLHSMATYDPLGLCYLMEYSNVR	84
Db	22	LITFPANVIMAILGNLWMAVACDROQRKTKTYFIVSLAPADLVSVYVNAFAGIEL	81
Qy	85	AEHCWYFGEVFCIKHTSTDIMLSASIFHLSFISIDRYAV-CDPLRYKAMNLLVICM	143
Db	82	VQDIWYEGMFCLVITSDLVTLTTSIFHLCCISIDRYAALCQPLVYRNMPFLRIALM	141
Qy	144	IFTSWVPAVFAFGMIFELNFKAAEELY-----KYNHRCGCGSVFPSKISGVLTFTSTF	199
Db	142	LGGCWYIIPMFISFLPMOGANNIGLVDIIEKKRKNHNSNFCVYPMWKNKEPALTCSYAF	201
Qy	200	YIPGSIIMLCVYRIRYILAKQARLISDANOQLOIGLEMKGISQSK-----	245
Db	202	YIPFLMLVLAAYRIVATLAKENH-----QQIQM-LQAGATSESRPQADQHSHTMRT	253
Qy	246	ERKAVYITLIGVAVGLICMCPEPCTWADPLAHYIIPPLNDVLWIMGVNSITFNPMVYA	305
Db	254	ETKRAKTLICVINGCCFCFCMAFFPYNIVDPIDYTVBEKWYATLWLGYNISGLNPLYA	313
Qy	306	FFYPWFRKALRKMFL	319
Db	314	FLNKSFRRAFLIIL	327

RESULT 4

dopamine receptor-like protein D14 - Japanese pufferfish
C.Species: Fugu rubripes (Japanese pufferfish)
C.Date: 27-Oct-1995 #sequence_revision 27-Oct-1995 #text_change 09-Jul-2004
C.Accession: A56849
R.Macrae, A.D.; Brenner, S.
Genomics 25, 436-446, 1995
A.Title: Analysis of the dopamine receptor family in the compact genome of the puffer fi
A.Reference number: A56849; MUID:95308911; PMID:7789977
A.Accession: A56849
A.Status: preliminary; not compared with conceptual translation
A.Molecule type: DNA
A.Residues: 1-459 <MAC>
A.Cross-references: UNIROT:P53452; GB:X80174; NID:g1204089; PIDN:CMA56455.1; PID:g12040
C.Superfamily: vertebrate rhodopsin
;Keywords: neurotransmitter receptor

Query Match	28.3%;	Score 510.5;	DB 2;	Length 459;
Best Local Similarity	37.0%;	Pred. No. 2e-35;		
Matches 118;	Conservative 54;	Mismatches 110;	Indels 37;	Gaps 8;

[illegible]


```

Qy      252  ILGIYMGVGLKCMKPEPFCITWMDPELHK-----IIPEPLNDVLIMFGYLNSTFNP 301
Db      283  TLTSLIMGFPVCCMLPPLFLNCMVPPCDRSPGHPOAGLPCVSETTFDIFVWFGMANSSLNP 342

Qy      302  MVTAFPFYPMFRKALKMML 319
Db      343  IIVAFNAD-FRKVFSSLL 359

RESULT 11
151659
dopamine D1A receptor - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: 151659
R:Sugamori, K.S.; Demchyshyn, L.L.; Chung, M.H.; Niznik, H.B.
Proc. Natl. Acad. Sci. U.S.A. 91, 10536-10540, 1994
A:Title: D1A, D1B, and D1C dopamine receptors from Xenopus laevis.
A:Reference number: 151659; PMID:7937989
A:Accession: 151659
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-451 <SUG>
A:Cross-references: UNIPROT:P42289; EMBL:U07863; NID:G559759; PIDD:AAA50828.1; PID:G559759
C:Superfamily: vertebrate rhodopsin
/Keywords: neurotransmitter receptor

```

```

Query Match      26.6%; Score 480; DB 2; Length 451;
Best Local Similarity 34.5%; Pred. No. 7.2e-33;
Matches 111; Conservative 56; Mismatches 117; Indels 38; Gaps 7;

Oy 32 LIILTLVGNLIIVSIVSHFQKLTNP-TNNLIHSMATVDPLGGLVMPYSWVSASHCQWY
Db 31 VLIISTLIGNTLVCAAVIRFRHLRSKVTNPFVIVISLANSDDLVAIVLWMPKVAALTAGWP 90
Oy 91 FGEVFECKIHSTDIIMLSASIFHLSPFISIDRYAVACPLRYKAKONTILVICMIFISWS 150
Db 91 FG-TFCNIWVAFDIMGCSPTASILNLCVISVDRYVAIISPFYERKOTPRKAVIRIMIGVAVTL 149
Oy 151 PAVAFAGCIFIELNPFKAEELIY-----KAVHCRGGCVSPFSKISGLVTMTSTFIYGS 204
Db 150 SVLISF--IPQLWMMHKRTTSFPDLNITLHRTMDNDSLSLRYAIISSLSFIYIPVA 207
Oy 205 IMLCVYRIYVIAKEOARLISDANOKLQIGLEMKNGISQ-----SKER 247
Db 208 IMITTYIRIYVIAAKQIRISALERAAVHAKGCONSTNRSLDCCQPESSLKTSFKRET 267
Oy 248 KAVKTLGIWGVPLICWCPPFICTVMDFFLH-----YIIPPLNDVLIWFGYLS 297
Db 268 KVLKTLISVIMGVPCWMLPFPLINCIYVFCPSPLTSTGEPEFCISSTTFDVFVFGWANS 327
Oy 298 TFENMVYVAFPFYMPRKALKMML 319
Db 328 SLNPIYAFNAD-PRKAFSNLL 348

RESULT 12
S688423
serotonin receptor 1D alpha - rabbit
N/Alternate names: 5-hydroxytryptamine receptor 1D alpha (5-HT1Dalpha)
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C/Accession: S688423
R/Hawood, G.; Lockyer, M.; Giles, H.; Fairweather, N.
FEBS Lett. 377, 73-76, 1995
A/Title: Cloning and characterization of the rabbit 5-HT(1D-alpha) and 5-HT(1D-beta) rec
A/Reference number: S688423; MUID:96130324; PMID:8545023
A/Accession: S688423
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-377 <HAR>
A/Cross-References: UNIPROT:P49145; EMBL:Z50162; NID:g1004281; PIDD:CA90530.1;
C/Superfamily: octopamine receptor type I

```

C:Keywords: neurotransmitter receptor

Query Match	26.3%	Score 475;	DB 2;	length 377;
Best local Similarity	32.0%;	Pred. No. 1.6e-32;		
Matches 112;	Conservative 62;	Mismatches 120;	Indels 56;	Gaps 7;
QY	22	VRASTLSLMLVILITTLVGNLIVVSLSHFQCHTPTNNMLHSMATVDFLLGCLVMPYSM	81	
Db	37	LKISLAVVLSITVATVLSNTEFVTLITLTKRHTPAVYLTGLSLATVTDLVSILVMPISI	96	
QY	82	VRSAEHCQFGEVCEVCKIHTSTDIMLSSASIFHLSPISIDRYAVCDPLRYAKONILVIC	141	
Db	97	AYTITHNNQGVLCIDIMWSSDIDCCNASIHLCLVIMLDRWALTDLMEYSKRTPAGHAA	156	
QY	142	VMITISVVPVAVPFQGMIFLELNFKGABEITYKXKHHVCHGCGSVFFSKS-GVLTFTMSFY	200	
Db	157	AMIVVVAISICISIPPLFWR-QAKAHBEV-----SDCLVNTSQISYTIYSTCGAFY	207	
QY	201	IPGSIIMCVVYRIYLIKEQ-----ARLISDA-----	227	
Db	208	IPSVLLIVLGRITVMAANRRLINPPLSYGKRTTAHLITSGAGSLCSLSLGBGSHS	267	
QY	228	-----NQKLIQGLEMKNGISOSKERRKAVXTLGIWGVPLICWCPPEFICVTMBPL	277	
Db	268	AGSPLFENPVARIKLADSVLKERKRISAAREKATVTLGIILGAFICWMLPEFVASLVLEIC	327	
QY	278	H-YIIPPTLNDVLIMFGYLNSTNPNMYAIFYMFRKALKMML-FGKIIF	324	
Db	328	RDSQCMPEPGLEDFETFWLGYLNSLNPJIYIVFNEDFRQAFQVRIHFRAKF	377	

[illegible]

```
QY 142 VNIISWSVPVAFPMIFELNPKGAEIYYKHVHCRGGCVFYSKIS-GVLTFTSPY 200
Db 157 TMAIYMAISICISIPFLFMR-QAKQDEM-----SDCLVNTSQISYTIYSTGARY 207
QY 201 IREGIMLCVYRIYLIJAKQ-----ARLISDA-----NOKIQIG----- 234
Db 208 IPSVLLIILYGRIRYRAARRIINPRLYGRKFTTAHLITGSAGSSILCSINSLHEGSHS 267
QY 235 -----LEMKGISQSKERKAVKTLGIVMGVFLICMCPFFICTYMDP- 275
Db 268 AGSPLEFNNVKTKLADSALEKRR-ISAAREKATKILGILGAFITCMLPFPVSLVLP 326
QY 276 -----FLAYIIPPTLNDVLWFGYLNSTFNPMVYAFYFWMFKALMMLFGKI--FOKDS 328
Db 327 CRDSCWIIH-----PALDFPFTWLGYLNSLNPILITVFNSEFQA-----FOKIVPRKAS 377

RESULT 14
I77467
serotonin receptor 1D - rat
N:Alternate names: 5-hydroxytryptamine receptor 1D (5-HT1D)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #ext_change 09-Jul-2004
C:Accession: I77467
R:Hamblin, M.W.; McGuffin, R.W.; Metcalf, M.A.; Dorsa, D.M.; Merchant, K.M.
Mol. Cell. Neurosci. 3, 578-587, 1992
A:Title: Distinct 5-HT1B and 5-HT1D serotonin receptors in rat: structural and pharmacol
A:Reference number: 157683
A:Accession: I77467
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-374 <RES>
A:Cross-references: UNIPROT:P28565; GB:M89953; NID:G202544; P1DN:AAA0614.1; PID:G202545
C:Superfamily: octopamine receptor type I
C:Keywords: neurotransmitter receptor

Query Match 26.2%; Score 473; DB 2; Length 374;
Best Local Similarity 32.1%; Pred. No. 2.3e-32;
Matches 117; Conservative 62; Mismatches 116; Indels 70; Gaps 10;

QY 18 MSNDV-----RASLSLMLVLIITLVGNLIYVSIHSFKQLHTPTNWLHSMATYDFLLG 73
Db 26 WDPEVLQALRIISLVVLSITLATVLSNAFVLTITLTKKLTPTNPLYIGSLATDDLVS 85
QY 74 CLVMPYSWRSAEHCWYFGEVFCIKHTSDIMLSASIFHLISIDRYAVCDPLRYKA 133
Db 86 ILVMPISATYTTTRTNFGQICDIWSSDITCTASILHLCVLADRYMALTDLEYSK 145
QY 134 KNNIIVICWMIPTISVPAVAFPMIFELNPKGAEIYYKHVHCRGGCVFYSKIS-GV 192
Db 146 RRTAGHAAMIAAVALISICISIPFLFMR-QATAHEM-----SDCLVNTSQISYTI 196
QY 193 LTFMSEYIPGIMLCVYRIYLIJAKQ-----ARLISDA----- 227
Db 197 YSTGAFYIPISLIIILYGRIRYRAARRIINPRLYGRKFTTAHLITGSAGSSILCSINSLNPS 256
QY 228 -----NQ-----KLQIGLEMKNGISQSKERKAVTIGIVMGVFLICMCPFI 269
Db 257 LHESHTVSGPLPFNNQKILADSIILERKRIISAAREKATKILGILGAFITCMLPFPV 316
QY 270 CTVNDP-----FLHYIIPPTLNDVLWFGYLNSTFNPMVYAFYFWMFKALMMLFGKI 323
Db 317 VSLVPLICRDSWIIH-----PALDFPFTWLGYLNSLNPILITVFNSEFQA-----FOKIVPRKAS 377

QY 324 FOKDS 328
Db 370 FRKAS 374

RESULT 15
JH0449
histamine H2 receptor - human
```

```
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #ext_change 09-Jul-2004
C:Accession: JH0449; I52319
R:Ganz, I.; Munzer, G.; Taehiro, T.; Schaeffer, M.; Wang, L.; DelValle, J.; Yamada, T
Biochem. Biophys. Res. Commun. 178, 1386-1392, 1991
A:Title: Molecular cloning of the human histamine H2 receptor.
A:Reference number: JH0449; MUID:91337087; PMID:1714721
A:Accession: JH0449
A:Molecule type: DNA
A:Residues: 1-359 <GAN>
A:Cross-references: UNIPROT:P25021; GB:M64799; NID:G184087; P1DN:AAA58647.1; PID:G184088
R:Nishi, T.; Koike, T.; Oka, T.; Maeda, M.; Futai, M.
Biochem. Biophys. Res. Commun. 210, 616-623, 1995
A:Title: Identification of the promoter region of the human histamine H2-receptor gene.
A:Reference number: I52319; MUID:95275318; PMID:7755641
A:Accession: I52319
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-359 <RES>
A:Cross-references: GB:P49783; NID:G728495; P1DN:BAA0818.1; PID:G728496
C:Comment: Histamine is important in the regulation of gastric acid secretion.
C:Genetics:
A:Gene: GDB:HRH2
A:Map position: 5pter-5qter
C:Superfamily: G protein-coupled receptor; transmembrane protein
C:Keywords: G protein-coupled receptor; transmembrane protein
F:23-44/Domain: transmembrane #status predicted <TM1>
F:58-81/Domain: transmembrane #status predicted <TM2>
F:93-114/Domain: transmembrane #status predicted <TM3>
F:135-159/Domain: transmembrane #status predicted <TM4>
F:181-204/Domain: transmembrane #status predicted <TM5>
F:235-258/Domain: transmembrane #status predicted <TM6>
F:268-289/Domain: transmembrane #status predicted <TM7>

Query Match 26.2%; Score 472.5; DB 2; Length 359;
Best Local Similarity 32.6%; Pred. No. 2.5e-32;
Matches 106; Conservative 64; Mismatches 124; Indels 31; Gaps 5;

QY 19 SNDVRAISLMLVLIITLVGNLIYVSIHSFKQLHTPTNWLHSMATYDFLLGCLVMP 78
Db 14 STACKITTVLAVLILITVAGNVVCLAVGLNRRNLNCFVLSLATIDLLGLVLP 73
QY 79 YSMWRSAEHCWYFGEVFCIKHTSDIMLSASIFHLISIDRYAVCDPLRYKKNIL 138
Db 74 FSAIYQISCKRSFKVCNITYISDVLCCTASILNLMISLDKCAVMDPLRYVVLTPV 133
QY 139 VICWMIPTISVPAVAFPMIFELNPKGAEIYYKHVHCRGGCVFYSKISGVLTFTMS 198
Db 134 RVALSLVLIWISITLSFSLIHLGMSNRNETS--KGNHTTSKCKVQVNEVYGLVDGLVT 190
QY 199 FYIPGSIIMLCVYRIYLIJAKQARLISDANOKIQIGLEMKNGISQSK-----ERKAVKT 253
Db 191 FYIPLIMCTTYRIIFVARDQAKRI-----NHISWGAATIRERKAVTL 236
QY 254 GIWVGVLICMCPFFICTVW-----DPLHYIIPPTLNDVLWFGYLNSTFNPMVYAFY 308
Db 237 AAWVGAFIILCMFPFTFYVRGLRGDAINEV-----LEALVNLGYANGSLNPILYALN 292
QY 309 PWRFKALMMLFGKIFQKSSRCKL 333
Db 293 RDRFTGYQQLFCRLANRNSHKTSL 317

Search completed: February 12, 2005, 03:45:11
Job time : 23 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 02:35:58 ; Search time 72 Seconds
(without alignments)
2411.040 Million cell updates/sec

Title: US-09-980-145-6
Perfect score: 1806
Sequence: 1 MPPCHNINISCVNMSN.....FGKIFQKDSRCKLFLRLSS 339

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1806	100.0	339	1	TAR1_HUMAN
2	1757	97.3	338	2	TAR1_MACMU
3	1403	77.7	332	1	TAR1_MOUSE
4	1351	74.8	332	1	TAR1_MOUSE
5	927.5	51.4	347	1	TAR2_RAT
6	817	45.2	306	2	O6NWS2
7	817	45.2	306	2	O6NWS3
8	815	45.1	306	1	GP58_HUMAN
9	815	45.1	306	2	O6NWS1
10	813.5	45.0	343	1	GP57_HUMAN
11	782	43.3	338	1	TAR3_RAT
12	743	41.1	348	1	TAR3_HUMAN
13	741	41.0	333	1	TAR12_RAT
14	733	40.6	344	1	TAR7_RAT
15	733	40.6	345	1	TAR4_RAT
16	731	40.5	344	1	TAR11_RAT
17	730	40.4	344	1	TAR10_RAT
18	722	40.0	358	1	TAR14_RAT
19	722	40.0	358	1	TAR15_RAT
20	715	39.6	358	1	TAR4_HUMAN
21	713	39.5	345	1	TAR8_RAT
22	710	39.3	358	1	TAR9_RAT
23	704	39.0	358	1	TAR6_RAT
24	697	38.6	342	1	TAR5_HUMAN
25	679	37.6	362	1	TAR13_RAT
26	677.5	37.5	337	2	O14804
27	668.5	37.0	337	2	O6NWT8
28	580.5	32.1	352	2	O9YHV4
29	545.5	30.2	328	2	O9YHV8
30	542.5	30.0	388	1	5H4_CAVPO
31	539	29.8	388	1	5H4_MOUSE

32	534	29.6	388	1	5H4_HUMAN	Q13639	homo sapien
33	534	29.6	388	2	O81XH9	Q81XH9	homo sapien
34	534	29.6	428	2	O712M9	O712M9	homo sapien
35	529	29.3	406	1	5H4_RAT	O62758	rattus norv
36	526.5	29.2	328	2	O9YHV7	O9YHV7	fugu rubrip
37	519	28.7	400	2	O6R1B8	O6R1B8	sus scrofa
38	519	28.7	402	2	O6Q253	O6Q253	sus scrofa
39	511	28.3	353	2	O9YHV3	O9YHV3	lancepeta fl
40	510.5	28.3	459	1	D1DR_FUGRU	P53452	fugu rubrip
41	503	27.9	446	2	O42315	O42315	cyprinus ca
42	498.5	27.6	445	2	O98842	O98842	anguilla an
43	497.5	27.5	465	1	DCDR_XENIA	P42291	xenopus lae
44	495	27.4	437	2	O42316	O42316	cyprinus ca
45	493.5	27.3	363	1	D1DR_CARAU	P35406	carassius a

ALIGNMENTS

RESULT 1
TAR1_HUMAN STANDARD: PRT: 339 AA.
ID TAR1_HUMAN
AC Q9ERU0;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 1 (Tar-1).
GN Name=TAR1; Synonyms=TAL, TAR1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;
RA Borowsky B., Adam N., Jones K.A., Raddeatz R., Artyushyn R.,
RA Ogozalek K.L., Durkin M.M., Lakhiani P.P., Bonini J.A., Pathirana S.,
RA Boyle N., Pu X., Kouranova E., Lichtblau H., Ochoa F.Y.,
RA Branchek T.A., Gerald C.;
RT "Trace amines: identification of a family of mammalian G protein-
coupled receptors.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21580235; PubMed=11723224;
RA Bunzow J.R., Sonders M.S., Artamangkul S., Harrison L.M., Zhang G.,
RA Quigley D.I., Darland T., Suchland K.L., Pasumamula S., Kennedy J.L.,
RA Olsson S.B., Magenlis R.E., Amara S.G., Grandy D.K.;
RT "Amphetamine, 3,4-methylenedioxymethamphetamine, lysergic acid
diethylamide, and metabolites of the catecholamine neurotransmitters
are agonists of a rat trace amine receptor.";
RL Mol. Pharmacol. 60:1181-1188(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Kopetz S.A., Aronstam R.S., Sharma S.V.;
RT "cDNA clones of human proteins involved in signal transduction
stimulated by the G-protein coupled receptor (www.cdna.org).";
RL Submitted (Nov-2002) to the EMBL/Genbank/DBJ databases.
-1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.
-1- Activated by endogenous trace amines as well as metabolites of the
biogenic amine neurotransmitters. Trace amines are biogenic amines
present in very low levels in mammalian tissues. Although some
trace amines have clearly defined roles as neurotransmitters in
invertebrates, the extent to which they function as true
neurotransmitters in vertebrates has remained speculative. Trace
amines are likely to be involved in a variety of physiological
functions that have yet to be fully understood. This receptor
seems to be mediated by the G(s)-class of G-proteins which
activate adenylate cyclase.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- TISSUE SPECIFICITY: Detected in low levels in discrete regions
within the central nervous system and in several peripheral
tissues. Moderately expressed in stomach. Low levels in amygdala,

CC kidney, and lung, and small intestine. Trace amounts in cerebellum,
CC dorsal root ganglia, hippocampus, hypothalamus, liver, medulla,
CC pancreas, pituitary, pontine reticular formation, prostate,
CC skeletal muscle, and spleen.

-1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF380185; AAK71236.1; -

DR EMBL; AF200627; AAG17112.1; -

DR EMBL; AY180374; AAO22154.1; -

DR Genew; HGNC:17734; TRAR1.

DR InterPro; IPR000276; GPCR_Rhodopsin.

DR InterPro; IPR009132; Tracamine_recept.

DR InterPro; IPR009133; Tracamine_recept1.

DR Pfam; PF00001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHHODOPSIN.

DR PRINTS; PR01831; TRACAMINER.

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.

DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.

DR G-protein coupled receptor; Glycoprotein; Multigene family;

KW Transmembrane.

FT DOMAIN 1 25

FT TRANSMEM 26 46

FT DOMAIN 47 59

FT TRANSMEM 60 80

FT DOMAIN 81 98

FT TRANSMEM 99 119

FT DOMAIN 120 136

FT TRANSMEM 137 157

FT DOMAIN 158 188

FT TRANSMEM 189 209

FT DOMAIN 210 252

FT TRANSMEM 253 273

FT DOMAIN 274 287

FT TRANSMEM 288 308

FT DOMAIN 309 339

FT TRANSMEM 340 382

FT DISULFID 39 10

FT CARBOHYD 10 10

FT CARBOHYD 17 17

SEQUENCE 339 AA; 39091 MW; 5E72FA61CFAC0E0 CRC64;

Query Match 100.0%; Score 1806; DB 1; Length 339;

Best Local Similarity 100.0%; Pred. No. 3.8e-117;

Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPFGHNIINISCVKNNNSNDVRASLYSLMWLIIITLVGNLIYVISISHRKQLATPTNW 60
DB 1 MNPFGHNIINISCVKNNNSNDVRASLYSLMWLIIITLVGNLIYVISISHRKQLATPTNW 60
QY 61 LIHSAATYDFLLGLCWMPYSWRSASHEQWGEVCKTHSTDTMLMSASIFHLSPISID 120
DB 61 LIHSAATYDFLLGLCWMPYSWRSASHEQWGEVCKTHSTDTMLMSASIFHLSPISID 120
QY 121 RYVAVCDPLRYKAKNNILVICMIFISWSVPAVAFGMIFELNFKGAEIYYGHVCRG 180
DB 121 RYVAVCDPLRYKAKNNILVICMIFISWSVPAVAFGMIFELNFKGAEIYYGHVCRG 180
QY 181 GCSVFEKISGVLFEMTSFYTPGSIMLCVYRRIYLIAKEQARLISDANQKQIGLEWKG 240
DB 181 GCSVFEKISGVLFEMTSFYTPGSIMLCVYRRIYLIAKEQARLISDANQKQIGLEWKG 240
QY 241 ISGSKERKAVKTLGIWGVFLICWCPFFICTVMDPELHYIIPPLNDVYLWPGVNSTEN 300
DB 241 ISGSKERKAVKTLGIWGVFLICWCPFFICTVMDPELHYIIPPLNDVYLWPGVNSTEN 300

QY 301 PMVYAFYFWFKALKMLFGKIFQKDSRRCLFLELSS 339
DB 301 PMVYAFYFWFKALKMLFGKIFQKDSRRCLFLELSS 339

RESULT 2

TABLE 1
ID TARI_MACMU STANDARD; PRT; 338 AA.
AC Q8H264;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 1 (Tar-1).
GN Name=TRAR1; Synonyms=TAR1;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]

RP SEQUENCE FROM N.A.

RA Miller G.M., Madras B.K.;

RT "Cloning of trace amine receptor 1 (TAR1) from Rhesus monkey.";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.

biogenic amine neurotransmitters (by similarity). Trace amines are

biogenic amines present in very low levels in mammalian tissues.

Although some trace amines have clearly defined roles as

neurotransmitters in invertebrates, the extent to which they

function as true neurotransmitters in vertebrates has remained

speculative. Trace amines are likely to be involved in a variety

of physiological functions that have yet to be fully understood.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; AY153366; AAN06172.1; -

InterPro; IPR000276; GPCR_Rhodopsin.

InterPro; IPR009132; Tracamine_recept.

InterPro; IPR009133; Tracamine_recept1.

Pfam; PF00001; 7tm_1; 1.

PRINTS; PR00237; GPCRHHODOPSIN.

PRINTS; PR01831; TRACAMINER.

PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.

PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.

G-protein coupled receptor; Glycoprotein; Multigene family;

KW Transmembrane.

FT DOMAIN 1 24

FT TRANSMEM 25 45

FT DOMAIN 46 58

FT TRANSMEM 59 79

FT DOMAIN 80 97

FT TRANSMEM 98 118

FT DOMAIN 119 135

FT TRANSMEM 136 156

FT DOMAIN 157 187

FT TRANSMEM 188 208

FT DOMAIN 209 251

FT TRANSMEM 252 272

FT DOMAIN 273 287

FT TRANSMEM 287 307

FT DOMAIN 308 338

FT DISULFID 95 181

FT CARBOHYD 9 9

N-linked (GlcNAc. . .) (Potential).

FT CARBOHYD 16 16 N-linked (GlcNAc...) (potential).
 SQ SEQUENCE 338 AA; 38797 MW; 98E7B35456B9409B CRC64;
 Query Match 97.3%; Score 1757; DB 1; Length 338;
 Best Local Similarity 96.4%; Pred. No. 9,3e-114;
 Matches 326; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 MPFCNINISCVKNNKNDVRAISYSLMVLIIITLTVGNLIVISISHFKQLHPTPTWL 61
 DB 1 MPFCNININISCVKNNKNDVRAISYSLMALIITLTVGNLIVISISHFKQLHPTPTWL 60
 QY 62 IHSNATVDFLLGCLVMPYSWVSAEHCWYFGFVCKHTSTDMILSSAIFHLSPFISIDR 121
 DB 61 IHSNATVDFLLGCLVMPYSWVSAEHCWYFGFVCKHTSTDMILSSAIFHLSPFISIDR 120
 QY 122 YVAVCDPLRYAKKNILVICWIFISWSVPAVAFGMIFLEINFGABEIIYKVVHCRGG 181
 DB 121 YVAVCDPLRYAKKNILVICWIFISWSVPAVAFGMIFLEINFGABEIIYKVVHCRGG 180
 QY 182 CSVFSPKISGVLTFTSFYIPGSIIMCYVYRIYLAKQARLISDANQKLGLEMKNGI 241
 DB 181 CSVFSPKISGVLTFTSFYIPGSIIMCYVYRIYLAKQARLISDANQKLGLEMKNGI 240
 QY 242 SOSKERKAVKTLGIWGVFLICWCPFFICTVMDPELHYIIPPTLNDVLIWFGYLNSTNP 301
 DB 241 SOSKERKAVKTLGIWGVFLICWCPFFICTVMDPELHYIIPPTLNDVLIWFGYLNSTNP 300
 QY 302 MYVAFYPMFRKALMFLGKIFORDSSRCKLFLELSS 339
 DB 301 MYVAFYPMFRKALMFLGKIFORDSSRCKLFLELSS 338

RESULT 3
 TART1 RAT STANDARD; PRT; 332 AA.
 AC 0923Y9; Q8VH05;
 DT 29-MAR-2004 (Rel. 43; Created)
 DT 29-MAR-2004 (Rel. 43; Last sequence update)
 DT 25-OCT-2004 (Rel. 45; Last annotation update)
 DE Trace amine receptor 1 (Tart1).
 GN Name=Tart1; Synonym=Ta1, TART1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT GLN-170.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;
 RA Borowsky B., Adam N., Jones K.A., Raddatz R., Artymshyn R.,
 RA Ogorzalek K.L., Durkin M.M., Lakhani P.P., Bonini J.A., Pathirana S.,
 RA Boyle N., Pu X., Kouranova E., Lichtblau H., Ochoa F.Y.,
 RA Branchek T.A., Gerald C.;
 RT "Trace amines: identification of a family of mammalian G protein-
 RT coupled receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Cerebellum, and Pancreatic tumor;
 RX MEDLINE=21580235; PubMed=11723224;
 RA Bunzow J.R., Sonders M.S., Artamangkul S., Harrison L.M., Zhang G.,
 RA Quigley D.I., Darland T., Suchland K.L., Pasumamula S., Kennedy J.L.,
 RA Olson S.B., Megret R.E., Amara S.G., Grandy D.K.;
 RT "amphetamine, 3,4-methylenedioxymethamphetamine, lysergic acid
 RT diethylamide, and metabolites of the catecholamine neurotransmitters
 RT are agonists of a rat trace amine receptor.";
 RL Mol. Pharmacol. 60:1181-1188(2001).
 CC -1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.
 CC Activated by endogenous trace amines as well as metabolites of the
 CC biogenic amine neurotransmitters. Trace amines are biogenic amines
 CC present in very low levels in mammalian tissues. Although some
 CC trace amines have clearly defined roles as neurotransmitters in
 CC invertebrates, the extent to which they function as true

CC neurotransmitters in vertebrates has remained speculative. Trace
 CC amines are likely to be involved in a variety of physiological
 CC functions that have yet to be fully understood. This receptor
 CC seems to be mediated by the G(s)-class of G-proteins which
 CC activate adenylylate cyclase.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC TISSUE SPECIFICITY: Widely distributed, but in low abundance,
 CC throughout the brain. Highest levels detected in the olfactory
 CC bulb, nucleus accumbens/olfactory tubercle, prefrontal cortex and
 CC other cortical regions, midbrain regions consisting of substantia
 CC nigra and ventral tegmentum, cerebellum, and pons/medulla. Among
 CC peripheral tissues, highest level observed in the liver, less
 CC expression in kidney, gastrointestinal tract, spleen, pancreas,
 CC and heart.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF380186; AAK71237.1; -;
 CC DR EMBL; AF421352; AAL65137.1; -;
 CC DR RGD; 621621; Tart1.
 CC DR InterPro; IPR000276; GPCR_Rhodopsn.
 CC DR InterPro; IPR009132; Tracamine_recept.
 CC DR InterPro; IPR009133; Tracamine_receptl.
 CC DR Pfam; PF00001; 7tm_1; 1.
 CC DR PRINTS; PR00237; GPCRHOODPSN.
 CC DR PRINTS; PR01831; TRACAMINER.
 CC DR PRINTS; PR01830; TRACAMINER.
 CC DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 CC DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 CC KW G-protein coupled receptor; Glycoprotein; Multigene family;
 CC KM Polymorphism; Transmembrane.
 CC FT DOMAIN 1 24
 CC FT TRANSMEM 25 45
 CC FT DOMAIN 46 58
 CC FT TRANSMEM 59 79
 CC FT DOMAIN 80 97
 CC FT TRANSMEM 98 118
 CC FT DOMAIN 119 138
 CC FT TRANSMEM 139 159
 CC FT DOMAIN 160 187
 CC FT TRANSMEM 188 208
 CC FT DOMAIN 209 249
 CC FT TRANSMEM 250 270
 CC FT DOMAIN 271 287
 CC FT TRANSMEM 288 308
 CC FT DOMAIN 309 332
 CC FT DISULFID 95 181
 CC FT CARBOHYD 9 16
 CC FT CARBOHYD 16 16
 CC FT CARBOHYD 283 283
 CC FT VARIANT 170 170
 CC SQ SEQUENCE 332 AA; 38021 MW; AD7F3A728C77B246 CRC64;
 Query Match 77.7%; Score 1403; DB 1; Length 332;
 Best Local Similarity 78.1%; Pred. No. 2,7e-89;
 Matches 261; Conservative 27; Mismatches 44; Indels 2; Gaps 1;

QY 2 MPFCNINISCVKNNKNDVRAISYSLMVLIIITLTVGNLIVISISHFKQLHPTPTWL 61
 DB 1 MPFCNININISCVKNNKNDVRAISYSLMALIITLTVGNLIVISISHFKQLHPTPTWL 60
 QY 62 IHSNATVDFLLGCLVMPYSWVSAEHCWYFGFVCKHTSTDMILSSAIFHLSPFISIDR 121
 DB 61 IHSNATVDFLLGCLVMPYSWVSAEHCWYFGFVCKHTSTDMILSSAIFHLSPFISIDR 120
 QY 122 YVAVCDPLRYAKKNILVICWIFISWSVPAVAFGMIFLEINFGABEIIYKVVHCRGG 181

```

Db 121 YVAVCDPLRYKAKINLAIPIVILISMSLPAVAFPGMIFLELNAGVLELHYNOVFLRG 180
QY 182 CSVFESKISGVLTMTSTFYIPGSMLCVYRYIYIAEQARLISDANOQLQIGLEMKNGI 241
Db 181 CFFPFSKVSGLVAMTSTFYIPGSMVLFYRYIYIAKQARSINRAN--LVQGLEGSRA 238
QY 242 SOSKERKAVKTLGIWGVFLICMCPEFLCTVMDPELHYIIPPTLNDVLIWPGYLNSTNP 301
Db 239 POSKETKAKTLGIWGVFLICMCPEFLCTVMDPELHYIIPPTLNDVLIWPGYLNSTNP 298
QY 302 MYAFYFPWFRALKKMLFGKIFOKDSRSCKLFL 335
Db 299 MYAFYFPWFRALKKMLFGKIFOKDSRSCKLFL 332

```

RESULT 4

```

TAR1_MOUSE STANDARD; PRT; 332 AA.

```

```

AC Q923Y8;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 1 (Tar1).
GN Name=Trar1; Synonyms=Tal, Tar1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;
RA Borowsky B., Adham N., Jones K.A., Raddatz R., Artymushyn R.,
RA Ogozalek K.L., Durkin M.W., Lakhiani P.P., Bonini J.A., Pathirana S.,
RA Boyle N., Pu X., Kouranova E., Lichtblau H., Octoa F.Y.,
RA Branchek T.A., Gerald C.;
"Trace amines: identification of a family of mammalian G protein-
RT coupled receptors.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).

```

```

-1- FUNCTION: Orphan receptor. Could be a receptor for trace amine.
    Activated by endogenous trace amines as well as metabolites of the
    biogenic amine neurotransmitters (By similarity). Trace amines are
    biogenic amines present in very low levels in mammalian tissues.
    Although some trace amines have clearly defined roles as
    neurotransmitters in invertebrates, the extent to which they
    function as true neurotransmitters in vertebrates has remained
    speculative. Trace amines are likely to be involved in a variety
    of physiological functions that have yet to be fully understood.
    This receptor seems to be mediated by the G(s)-class of G-proteins
    which activate adenylate cyclase (By similarity).
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- TISSUE SPECIFICITY: Widely distributed throughout the brain.
    Strongly expressed in the mitral cell layer of the olfactory bulb,
    piriform cortex, the arcuate, motor, and mesencephalic trigeminal
    nuclei, lateral reticular and hypoglossal nuclei, cerebellar
    Purkinje cells, and ventral horn of the spinal cord. Moderately
    expressed in the frontal, entorhinal, and agranular cortices, the
    ventral pallidum, thalamus, hippocampus, several hypothalamic
    nuclei, amygdala, dorsal raphe, and gigantocellular reticular
    nuclei. Weakly expressed in the septum, basal ganglia, amygdala,
    myelencephalon, and spinal cord dorsal horn. Particularly
    interesting is the moderate expression in several monoaminergic
    cell groups, namely the dorsal raphe, the locus coeruleus, and the
    ventral tegmental area.
-1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1ab-sib.ch/announce/

```

or send an email to license@sib-sib.ch).

```

CC EMBL; AF380187; AK71238.1; -.
DR MGD; MG1:2148258; Tar1.
DR GO; GO:0016021; C: integral to membrane, IC.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR009132; Tracamine_recept.
DR InterPro; IPR009133; Tracamine_recept.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR01831; TRACAMINER.
DR PRINTS; PR01830; TRACAMINER.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS50262; G-PROTEIN RECEPTOR FL 2; 1.
KW G-protein coupled receptor; Glycoprotein; Multigene family;
KW Transmembrane.
FT DOMAIN 1 24 Extracellular (Potential).
FT TRANSMEM 25 45 1 (Potential).
FT DOMAIN 46 58 Cytoplasmic (Potential).
FT TRANSMEM 59 79 2 (Potential).
FT DOMAIN 80 97 Extracellular (Potential).
FT TRANSMEM 98 118 3 (Potential).
FT DOMAIN 119 138 Cytoplasmic (Potential).
FT TRANSMEM 139 159 4 (Potential).
FT DOMAIN 160 187 Extracellular (Potential).
FT TRANSMEM 188 208 5 (Potential).
FT DOMAIN 209 249 Cytoplasmic (Potential).
FT TRANSMEM 250 270 6 (Potential).
FT DOMAIN 271 287 Extracellular (Potential).
FT TRANSMEM 288 308 7 (Potential).
FT DOMAIN 309 332 Cytoplasmic (Potential).
FT DISULFID 95 181 By similarity.
FT CARBOHYD 9 9 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 332 AA; 37620 MW; 680B62B1892264 CRC64;

```

Query Match 74.8%; Score 1351; DB 1; Length 332;
 Best Local Similarity 75.4%; Pred. No. 1e-85;
 Matches 252; Conservative 31; Mismatches 49; Indels 2; Gaps 1;

```

QY 2 MPFENITINISCVKNNMSNDVRSALYSMLVILITLTVGNLVIYVSIHSKQLHTPTNML 61
Db 1 MHCHALTNISNNSWMSREVQASLYSLMSLITLITVGNLVIYVSIHSKQLHTPTNML 60
QY 62 IHSMAVDFPLGLVNPYSWVRSAGWYEGVFCKIHTSTDMSSASIFHSFISIDR 121
Db 61 LHSMAVDFPLGLVNPYSWVRSAGWYEGVFCKIHTSTDMSSASIFHSFISIDR 120
QY 122 YVAVCDPLRYKAKINLAIPIVILISMSLPAVAFPGMIFLELNAGVLELHYNOVFLRG 181
Db 121 YVAVCDPLRYKAKINLAIPIVILISMSLPAVAFPGMIFLELNAGVLELHYNOVFLRG 180
QY 182 CSVFESKISGVLTMTSTFYIPGSMLCVYRYIYIAEQARLISDANOQLQIGLEMKNGI 241
Db 181 CFFPFSKVSGLVAMTSTFYIPGSMVLFYRYIYIAKQARSINRAN--LVQGLEGSRA 238
QY 242 SOSKERKAVKTLGIWGVFLICMCPEFLCTVMDPELHYIIPPTLNDVLIWPGYLNSTNP 301
Db 239 POSKETKAKTLGIWGVFLICMCPEFLCTVMDPELHYIIPPTLNDVLIWPGYLNSTNP 298
QY 302 MYAFYFPWFRALKKMLFGKIFOKDSRSCKLFL 335
Db 299 MYAFYFPWFRALKKMLFGKIFOKDSRSCKLFL 332

```

RESULT 5

```

TAR2_MOUSE STANDARD; PRT; 347 AA.

```

```

AC Q923Y7;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 2 (Tar2).
GN Name=Trar2; Synonyms=Tar2, Tar2;

```



```

RA  Marchese A.;
RT  "Cloning and characterization of additional members of the G protein-
RU  coupled receptor family."
RU  Biochim. Biophys. Acta 1490:311-323(2000).
CC  -1- FUNCTION: Orphan receptor.
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC  -1- TISSUE SPECIFICITY: Not expressed in the pons, thalamus,
CC  hypothalamus, hippocampus, caudate, putamen, frontal cortex, basal
CC  forebrain, midbrain or liver.
CC  -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF112460; AAF27278.1; -.
DR  GenBank; HGNC:4514; GPR58.
DR  MIM; 604849; -.
DR  GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
DR  InterPro; IPR000276; GPCR_Rhodpsn.
DR  InterPro; IPR009132; Tracamine_recept.
DR  Pfam; PF00001; 7tm1.1; 1.
DR  PRINTS; PR00237; GPCR_RHODOPSIN.
DR  PRINTS; PR01830; TRACAMINER.
DR  PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR  PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW  G-protein coupled receptor; Glycoprotein; Transmembrane.
FT  DOMAIN 1 3 24 1 (Potential).
FT  TRANSMM 4 24 34 1 (Potential).
FT  TRANSMM 35 55 2 (Potential).
FT  TRANSMM 56 73 3 (Potential).
FT  TRANSMM 95 117 4 (Potential).
FT  TRANSMM 118 138 4 (Potential).
FT  TRANSMM 139 162 5 (Potential).
FT  TRANSMM 163 183 5 (Potential).
FT  TRANSMM 184 218 6 (Potential).
FT  TRANSMM 219 239 6 (Potential).
FT  TRANSMM 240 254 7 (Potential).
FT  TRANSMM 255 277 7 (Potential).
FT  TRANSMM 278 306 7 (Potential).
FT  DOMAIN 71 156 1 (Potential).
FT  DISULFD 244 244 1 (Potential).
FT  CARBOHYD 244 244 1 (Potential).
SQ  SEQUENCE 306 AA; 34924 MW; 55629F613062777C CRC64;

Query Match 45.1%; Score 815; DB 1; Length 306;
Best Local Similarity 50.0%; Pred. No. 1.1e-48;
Matches 158; Conservative 48; Mismatches 86; Indels 24; Gaps 4;

QY  LYSIMLTLLITLVGNLIVIVSISHFQKQHTPTNMLHISMATVDLGLCLVMPYSMTASA 85
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB  1 MYSPAGSIFITIFGNLMIISISYFKQKHTPTNLLISMATDPLGLFTIMPYSMISV 60
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  86 EHCWFGVGVFCKIKHTSTDIMLSSASIFHLSFISIDRYAVCDPLRYKAKMNLIVICWIF 145
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB  61 ENCWYFGTLTFCIKIYSPFLMLSTITFHCVAIRFAVACPLLYSTKITIPVKKRLIL 120
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  146 ISMSVPAVAFARGMIFLELNFGKAEIYYKVAHCRGGCVFSGISGVLTFFMISFYIPSI 205
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB  121 LCMWSVPAFAFGAVSEAVADIEG-YDILVACSSCPFMWKMLGTTLFMAGFTPSM 179
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  206 MLCVYRIYVLLAKEDARLISD--ANQKQIGLEMKNGISQSKERKAVKTLGITMGVPLIC 263
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB  180 MVGIGIKI FAVSRKAHAIINLRNNQNV-----KKDKAKATLIGIVIGVPLIC 229
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY  264 WCPPICTVMDPPLHIIIPITNDVLIWFGYNSTFNPMVTAFFYPMWRKALQMPLFGKI 323
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB  230 WPCPFITLDBPLNFSTPVVLFDAITWFGYNSTCNPLIYGFYPWFRALKYILGKI 289
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY  324 F-----QKDS 328
    |||:
DB  290 FSSCFNHTILCMOKES 305
    |||:

RESULT 9
ID  Q6NWS1 PRELIMINARY; PRT; 306 AA.
AC  Q6NWS1;
DT  05-JUL-2004 (TRENBLREL 27, Created)
DT  05-JUL-2004 (TRENBLREL 27, Last sequence update)
DT  05-JUL-2004 (TRENBLREL 27, Last annotation update)
DE  GPR58 protein (Fragment).
GN  Name=GPR58;
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N. A.
RC  TISSUE=PCR rescued clones;
RX  MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Aleesch S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA  Brownstein M.J., Ueda T.B., Tomshyuki S., Carrinci P., Prange C.,
RA  Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA  Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA  Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Skoloud J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA  Krzywnski M.I., Skalska U., Smallus D.E., Schermer A., Schein J.E.,
RA  Jones S.J., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RT  and mouse cDNA sequences."
RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN  [2]
RP  SEQUENCE FROM N. A.
RC  TISSUE=PCR rescued clones;
RA  Strausberg R.;
RA  Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC  -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
DR  EMBL; BC067463; AAH67463.1; -.
DR  GO; GO:0016021; C:Integral to membrane; IEA.
DR  GO; GO:0004872; F:receptor activity; IEA.
DR  GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR  GO; GO:0007186; P:G-protein coupled receptor protein signalin. . . IEA.
DR  InterPro; IPR000276; GPCR_Rhodpsn.
DR  InterPro; IPR009132; Tracamine_recept.
DR  Pfam; PF00001; 7tm1.1; 1.
DR  PRINTS; PR00237; GPCR_RHODOPSIN.
DR  PRINTS; PR01830; TRACAMINER.
DR  PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR  PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW  G-protein coupled receptor; Receptor; Transmembrane.
FT  NON_TER 306 306
SQ  SEQUENCE 306 AA; 34979 MW; 2D934A23607FFA03 CRC64;

Query Match 45.1%; Score 815; DB 2; Length 306;
Best Local Similarity 50.0%; Pred. No. 1.1e-48;
Matches 158; Conservative 49; Mismatches 85; Indels 24; Gaps 4;

QY  26 LYSIMLTLLITLVGNLIVIVSISHFQKQHTPTNMLHISMATVDLGLCLVMPYSMTASA 85
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB  1 MYSPAGSIFITIFGNLMIISISYFKQKHTPTNLLISMATDPLGLFTIMPYSMISV 60
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

QY 86 EHCWYFGEVCKIHTSTDMILSSASIFHLSFISIDRYAVCDPLRYKAKNNILVICMIF 145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ENCWYFGITFCCKIYYSDMLMSTISIFHLSVALDRYALCYLLSTKTTITVIRKLL 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 146 ISWVAVPAFAGMIFFLANKRGAELIYKAVHCRGGSVFPSKISGVLTMTSPYIPGSI 205
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 LCVSWPAGFAFAGVVFSEAYADGIEG-YDILVACSSSCPVMFNKLMGTTLFMAGFPTPGSM 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 206 MLCVYVRIYLAKQEARLISD--ANOKLOIGLEMKNGISOSKERKAVKTIGIMGVFLIC 263
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 MVGIYGIYFVSKRKAHAINNNKRNQNV-----KKQKKAKTIGIYIGVFLIC 229
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 264 WCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTFNPMVYAFYFPMFRALKMLFGKI 323
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 230 WFPCCFPIILDPLFNSTPVLVFDALKMGFGFNSTCNPLIYGFYFPMFRALKYIILGKI 289
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 324 F-----OKDS 328
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 290 FSSCFNHTILCWQES 305
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 10

GP57 HUMAN

ID GP57 HUMAN STANDARD; PRT; 343 AA.

```

AC Q9P1F4;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DN Probable G protein-coupled receptor GPR57.
GN Name=GPR57;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=20149852; PubMed=10684976; DOI=10.1016/S0167-4781(99)00241-9;
RA Lee D.K., Lynch K.R., Nguyen T., Im D.-S., Cheng R., Saldivia V.R.,
RA Liu Y., Liu I.S.C., Heng H.H.Q., Seeman P., George S.R., O'Dowd B.F.,
RA Marchese A.;
RT "Cloning and characterization of additional members of the G protein-
RT coupled receptor family."
RL Biochim. Biophys. Acta 1490:311-323 (2000).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Not expressed in the pons, thalamus, globus
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF112461; AAF27279.1; -.
DR Genew; HGNC:4513; GPR57.
DR MIM; 604848; -.
DR GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR009132; Tracamine_recept.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECPT_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECPT_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
FT DOMAIN 1 35 Extracellular (Potential).
FT TRANSMEM 36 56 1 (Potential).
FT DOMAIN 57 68 Cytoplasmic (Potential).

```

```

FT TRANSMEM 69 89 2 (Potential).
FT DOMAIN 90 150 Extracellular (Potential).
FT TRANSMEM 151 168 3 (Potential).
FT DOMAIN 169 172 Cytoplasmic (Potential).
FT TRANSMEM 173 193 4 (Potential).
FT DOMAIN 194 198 Extracellular (Potential).
FT TRANSMEM 199 223 5 (Potential).
FT DOMAIN 224 257 6 (Potential).
FT TRANSMEM 258 278 6 (Potential).
FT DOMAIN 279 287 Extracellular (Potential).
FT TRANSMEM 288 308 7 (Potential).
FT DOMAIN 309 343 Cytoplasmic (Potential).
FT CAROXYD 145 145 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 343 AA; 39065 MW; DCS1D15ABC026F CRC64;

```

Query Match 45.0%; Score 813.5; DB 1; Length 343;
 Best Local Similarity 46.0%; Pred. No. 1.5e-48;
 Matches 159; Conservative 57; Mismatches 99; Indels 31; Gaps 5;

```

QY 5 CHNIIN-----ISGVKN-----WSNDVRSLSYSLMVLITLTVLGNLIVYSISH 50
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 13 CPKFNKILSSHOPLFSCPDNVFVGYDSHDY-----PLFGNLVIMYSISH 58
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 51 FKQHTPTNMLIHSMAVDPLGLCLVMPYSVWSRSEHCWYFGEVCKIHTSTDMILSSAS 110
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 FKQHSPTNMLISMATTDPLGLGVIMPYSIMSVESCMVFGDGFCKFTSPDMRLTS 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 111 IFHLSFISIDRYAVCDPLRYKAKNNILVICMIFISWVAVPAFAGMIFLEINFKGAEE 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 IFHLSIAIDRFYAVCYPLHYITMTNSTIKQLALFACVSPALFSPFLVSEADVSGMOS 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 171 IYGHVHCRGGSVFPSKISGVLTMTSPFIRPSIMLCVYRYLYLAKQEARLISDANOK 230
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 -YKILVACFNPCALTFKFKNGITLFTCFPTPSIMWIGIKLFIYSKOHARVISHPEN 237
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 231 LQIGLEMKNGISOSKERKAVKTIGIYGVFLICWCFEICTVMDPLHYIIPPTLNDVLI 290
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 238 TKGA--VKHLSKKOKRKAKTIGIYGVFLACMLCFLAVLDPLVDYSTPIIIDLV 295
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 291 WFGYLNSTFNPMVYAFYFPMFRALKMLFGKIFQDSSRCKFLPE 336
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 296 WLRFPNSTCNPLIHGFENPWFQKAFYIVSGKIFSSHSETANLPE 341
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 11

TAR3 RAT

ID TAR3 RAT STANDARD; PRT; 338 AA.

```

AC Q923T6;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 3 (Tar-3).
GN Name=Trar3; Synonyms=Tr3; Tar3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP MEDLINE=15174364; PubMed=11459929; DOI=10.1073/pnas.151105198;
RA Borowsky B., Adham N., Jones K.A., Raddatz R., Artymehayn R.,
RA Gozalek K.L., Durkin M.M., Lakhani P.P., Bonini J.A., Pathirana S.,
RA Boyle N., Pu X., Kouranova E., Lichtblau H., Ochoa F.Y.,
RA Branchek T.A., Gerald C.;
RT "Trace amines: identification of a family of mammalian G protein-
RT coupled receptors."
RL Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971 (2001).
CC -!- FUNCTION: Orphan receptor. Could be a receptor for trace amines.
CC Trace amines are biogenic amines present in very low levels in
CC mammalian tissues. Although some trace amines have clearly defined
CC roles as neurotransmitters in invertebrates, the extent to which
CC they function as true neurotransmitters in vertebrates has

```


FT DOMAIN 169 197 Extracellular (Potential).
FT TRANSMEM 198 218 5 (Potential).
FT DOMAIN 219 259 Cytoplasmic (Potential).
FT TRANSMEM 260 280 6 (Potential).
FT DOMAIN 281 294 Extracellular (Potential).
FT TRANSMEM 295 315 7 (Potential).
FT DOMAIN 316 348 Cytoplasmic (Potential).
FT DISULFID 105 190 By similarity.
FT CARBOHYD 4 4 N-linked (GlcNAc..?) (Potential).
FT CARBOHYD 19 19 N-linked (GlcNAc..?) (Potential).
SQ SEQUENCE 348 AA, 39016 MW, DDC648480A6F761 CRC64;

Query Match 41.1%; Score 743; DB 1; Length 348;
Best Local Similarity 42.5%; Pred. No. 1.1e-43;
Matches 141; Conservative 63; Mismatches 126; Indels 2; Gaps 2;

QY 5 CHNINISGVKNMNDVRASISLMLVLLITLVGNLIYVSIHFKQLHTPTMNLHS 64
DB 14 CYKNVNESCTKTPYSPGPPSILLYAVLFGAVLAAGNLLMIAIDHFKQLHTPTNLAS 73
QY 65 MATVDFLLGLVMPYSWVSAHSCWYFGEVFCIKHTSTDIMLSASIFHLSPISIDRYA 124
DB 74 LACADFLVGVVMPSTYRSVSCWYFGDSYCKHTCPTSCFASLHLCISVDRIYA 133
QY 125 VCDPLRYAKKNILYICWMIFFISWVPAVFAFGMFLBNFGAEIYYKHVCRGGCV 184
DB 134 VTDPLTYPTKFTVSVSGICIVLWFFSVTSFSI FYTGANEGIBELVVA-LTCVGGCOA 192
QY 185 FFSKISGLTFTMTSPYIPGSIWLCVYVRYILAKQOALISANOKIQIGLEMKRIGSQS 244
DB 193 PLNQWVLLICFLVLPPIVAVAVFLYSKIFLVAKQAKRTISTASQAOSSSSYKERVAK 251
QY 245 KERKAVKTLGIWVGFLLICWCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTFNPVY 304
DB 252 RERKAKTLGLIAMAFLVSWLPYLDVAVIDAMNITTPRYVEILVWCYVNSANPLIY 311
QY 305 AFFYWPFRKALKMFLFKIFQKDSRCKLPLE 336
DB 312 AFFYWPFRKALKLIVSGKVLRTDSSTNLFSF 343

RESULT 13
TAR12_RAT 13
ID TAR12_RAT STANDARD; PRT; 333 AA.
AC Q923X8;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trace amine receptor 12 (Iar-12).
GN Name=Iar12; Synonyms=Iar12; Tar12;
OS Rattus norvegicus (Rat).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;
RA Borowczyk B., Adam N., Jones K.A., Raddatz R., Artymushyn R.,
RA Ogozalek K.L., Durkin M.W., Lakhiani P.P., Bonini U.A., Pathirana S.,
RA Boyle N., Pu X., Kouranova E., Lichtblau H., Ochoa F.Y.,
RA Branchek T.A., Gerald C.;
RT "Trace amines: identification of a family of mammalian G protein-
RT coupled receptors."
RT Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).
RL -1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.
CC Trace amines are biogenic amines present in very low levels in
CC mammalian tissues. Although some trace amines have clearly defined
CC roles as neurotransmitters in invertebrates, the extent to which
CC they function as true neurotransmitters in vertebrates has
CC remained speculative. Trace amines are likely to be involved in a
CC variety of physiological functions that have yet to be fully
CC understood.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).

CC EMBL: AF380200; AAK1251.1; -.
CC RGD: 631392; Tar12.
CC InterPro: IPR000276; GPCR_Rhodopsn.
CC InterPro: IPR009132; Tracamine_recept.
CC Pfam: PF00001; 7tm1_1.
CC PRINTS: PR00237; GPCR_HODOPSIN.
CC PRINTS: PR01830; TRACAMINER.
CC PROSITE: PS00237; G_PROTEIN_RECPT_F1_1.
CC PROSITE: PS50262; G_PROTEIN_RECPT_F1_2; 1.
CC G-protein coupled receptor; Glycoprotein; Multigene family;
KW Transmembrane.
FT DOMAIN 1 22 Extracellular (Potential).
FT TRANSMEM 23 43 1 (Potential).
FT DOMAIN 44 58 Cytoplasmic (Potential).
FT TRANSMEM 59 79 2 (Potential).
FT DOMAIN 80 99 Extracellular (Potential).
FT TRANSMEM 100 118 3 (Potential).
FT DOMAIN 119 141 Cytoplasmic (Potential).
FT TRANSMEM 142 162 4 (Potential).
FT DOMAIN 163 186 Extracellular (Potential).
FT TRANSMEM 187 207 5 (Potential).
FT DOMAIN 208 249 Cytoplasmic (Potential).
FT TRANSMEM 250 270 6 (Potential).
FT DOMAIN 271 283 Extracellular (Potential).
FT TRANSMEM 284 304 7 (Potential).
FT DOMAIN 305 333 Cytoplasmic (Potential).
FT DISULFID 95 180 By similarity.
FT CARBOHYD 9 N-linked (GlcNAc..?) (Potential).
SQ SEQUENCE 333 AA, 37467 MW, 65BD86131C1C95B CRC64;

Query Match 41.0%; Score 741; DB 1; Length 333;
Best Local Similarity 41.8%; Pred. No. 1.5e-43;
Matches 140; Conservative 70; Mismatches 123; Indels 2; Gaps 2;

QY 2 MPECHNINISGVKNMNDVRASISLMLVLLITLVGNLIYVSIHFKQLHTPTMNL 61
DB 1 MQLCYEKLNRSVNSPYSFGRLILYAVFGGAVLAVCGNLLVWTSILHFRQLHSNPVFL 60
QY 62 IHSNATVDFLLGLVMPYSWVSAHSCWYFGEVFCIKHTSTDIMLSASIFHLSPISIDR 121
DB 61 VASLACADFLVGLVMPYSWVSAHSCWYFGEVFCIKHTSTDIMLSASIFHLSPISIDR 120
QY 122 YVAVCDPLRYAKKNILYICWMIFFISWVPAVFAFGMFLBNFGAEIYYKHVCRGG 181
DB 121 YVAVSDPLRYAKKNILYICWMIFFISWVPAVFAFGMFLBNFGAEIYYKHVCRGG 179
QY 182 CSVFFSISGLTFTMTSPYIPGSIWLCVYVRYILAKQOALISANOKIQIGLEMKRIGI 241
DB 180 CQIADVNSWVPIINLLFLVPAVAVVYISKIFILAAQOAHNIBKMGKQPARASSEYKDR 238
QY 242 SOSKERAHVTLGIWVGFLLICWCPFFICTVMDPLHYIIPPTLNDVLIWFGYLNSTFNP 301
DB 239 VAKRERAAATLIGIAVAAPFLSLWPIYIDSIDAIFGVPVPTVYVELIHWIGYNSAMNP 298
QY 302 MTAFFYPWFRKALKMFLFKIFQKDSRCKLPLE 336
DB 299 LIAVAFYPWFRKALKLIVTKIIRENSATNLFSF 333

RESULT 14
TAR7_RAT 14
ID TAR7_RAT STANDARD; PRT; 344 AA.

Accession	Protein Name	Sequence	Score	DB 1	Length
Q923Y3	29-MAR-2004 (Rel. 43, Created)		40.6%	DB 1	Length 344
DT	29-MAR-2004 (Rel. 43, Last sequence update)		41.2%		
DT	25-OCT-2004 (Rel. 45, Last annotation update)				
DT	Trace amine receptor 7 (Tar-7)				
GN	Name=Trar7; Synonyms=Tar7, Tar7;				
OS	Rattus norvegicus (Rat)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxId=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Sprague-Dawley;				
RX	MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;				
RA	Botowary B., Adham N., Jones K.A., Raddatz R., Artyushyn R.,				
RA	Ogozalek K.L., Durkin M.M., Lakhani P.P., Bonini J.A., Pathirana S.,				
RA	Boyle N., Pu X., Kouranova E., Lichtblau H., Ochoa F.Y.,				
RA	Branchek T.A., Gerald C.;				
RT	"Trace amines: identification of a family of mammalian G protein-				
RT	coupled receptors."				
RL	Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).				
CC	-1- FUNCTION: Orphan receptor. Could be a receptor for trace amines.				
CC	Trace amines are biogenic amines present in very low levels in				
CC	mammalian tissues. Although some trace amines have clearly defined				
CC	roles as neurotransmitters in invertebrates, the extent to which				
CC	they function as true neurotransmitters in vertebrates has				
CC	remained speculative. Trace amines are likely to be involved in a				
CC	variety of physiological functions that have yet to be fully				
CC	understood.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	-1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; AF380195; AAK71246.1; -				
DR	RGD; 631386; Tar7				
DR	InterPro; IPR000276; GPCR_Rhodopsn				
DR	InterPro; IPR009132; Tracamine_recept.				
DR	Pfam; PF00001; 7tm_1; 1.				
DR	PRINTS; PR00237; GPCRHOODPSN.				
DR	PRINTS; PR01830; TRACEAMINER.				
DR	PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.				
DR	PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.				
KW	G-protein coupled receptor; Glycoprotein; Multigene family;				
KW	Transmembrane.				
FT	DOMAIN 1 33				
FT	TRANSMEM 34 54				
FT	DOMAIN 55 67				
FT	TRANSMEM 68 88				
FT	DOMAIN 89 102				
FT	TRANSMEM 103 127				
FT	DOMAIN 128 146				
FT	TRANSMEM 147 167				
FT	DOMAIN 168 196				
FT	TRANSMEM 197 217				
FT	DOMAIN 218 256				
FT	TRANSMEM 257 277				
FT	DOMAIN 278 295				
FT	TRANSMEM 296 319				
FT	DOMAIN 320 344				
FT	DISULFID 104 189				
FT	CARBOHYD 4 4				
FT	CARBOHYD 18 18				
FT	CARBOHYD 344 AA; 38043 MW; 45C41C8615F6408 CRC64;				
SQ	SEQUENCE				
Query Match	40.6%; DB 1; Length 344;				
Best Local Similarity	41.2%; Pred No. 5; 5e-43;				

Matches	139,	Conservative	73,	Mismatches	121,	Indels	4,	Gaps	3,
QY	2	MPFCHNIINISCVAKNMSNDVRASL	YSLAMLIIITTVGNLIVYSVSHRQHLTPTNNML	61					
Db	10	LQTCYENVNNAISCIPTPPSPGILRVLLVYVFGFGLAVACGNLVLIVSVLHPKQLHSPANFL		69					
QY	62	IHSATVDPFLGLMWPYSNVARSHEHCWGEVCKHTSTNDIMLSASIFHLSFISIDR	121						
Db	70	IASIASADPLVLGISVMPFSNWRSIESCWFYGDFFCSLHSCDDAAFCYSSLPHLCFISVDR	129						
QY	122	YVAACDPLRYAKAKNNILVICMWFISMSVAVAFAGMIPLFELNFKGAEIEIYKRVHRCGG	181						
Db	130	YIATPELVVPTKRTMSVSGICISISWILPLVYSANFYGISATGIENL-VSALNCVGG	188						
QY	182	CSVEFSKISGLVTFMTSPFYIPGSIIMLCVYRIYIYIAKEQARLI--SPANOKLQIGLEMKV	239						
Db	189	COVAINDDWVLISFL-FPIPLVMIILYSKIFLVAQCAVKIETSIISGSGESSLSHK	247						
QY	240	GISGSKRKAVKTIIGYMGVFLWCQCFPICTWDPDLPHVITPTLADVLIMFQYLNSTF	299						
Db	248	ARVAKRRKAKITGVTVMAFMWSMLPYTIDTLIDAMGFTTPAYVVEICGMAYNSAM	307						
QY	300	NPMVYAFYFPMFRKALKMMLFGKIFQDSSCKFLFLE	336						
Db	308	NPLIYAFYFPMFRKALKILISGLKILKXHSSTLSFSE	344						

RESULT 15

ID	TAR4_RAT	STANDARD;	PRT;	345 AA.
AC	Q923T5;			
DT	29-MAR-2004 (Rel. 43, Created)			
DT	29-MAR-2004 (Rel. 43, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Trace amine receptor 4 (Tar-4).			
GN	Name:Tar4; Synonyms:Ta4, Tar4;			
OS	Rattus norvegicus (Rat).			
OC	Eumetazoa; Chordata; Craniota; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RA	SEQUENCE FROM N.A.			
RP	STRAIN:Sprague-Dawley;			
RX	MEDLINE=21374364; PubMed=11459929; DOI=10.1073/pnas.151105198;			
RA	Boroweky B., Adam N., Jones K.A., Raddatz R., Artyomshyn R.,			
RA	Ogozalek K.L., Durkin M.M., Lakhani P.P., Bonini J.A., Pachlirana S.,			
RA	Boyle N., Pu X., Kouranova E., Iachubiau H., Ochoa F.Y.,			
RA	Branchek T.A., Gerald C.;			
RT	"Trace amines: identification of a family of mammalian G protein-			
RT	coupled receptors";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:8966-8971(2001).			
CC	-I- FUNCTION: Orphan receptor. Could be a receptor for trace amines.			
CC	Trace amines are biogenic amines present in very low levels in			
CC	mammalian tissues. Although some trace amines have clearly defined			
CC	roles as neurotransmitters in invertebrates, the extent to which			
CC	they function as true neurotransmitters in vertebrates has			
CC	remained speculative. Trace amines are likely to be involved in a			
CC	variety of physiological functions that have yet to be fully			
CC	understood.			
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-I- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; AF380191; AAK71242.1; -.			
DR	RGD; 631384; Ta4			
DR	InterPro; IPR000276; GPCR_Rhodpsn.			

DR InterPro; IPR009132; Tricamine_recept.
DR Pfam; PF00001; Tcm_1; 1.
DR PRINTS; PR00237; GPCR_HODOPSIN.
DR PRINTS; PR01830; TRACEAMINER.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Multigene family;
Transmembrane.
FT DOMAIN 1 32 Extracellular (Potential).
FT TRANSMEM 33 53 1 (Potential).
FT DOMAIN 54 68 Cytoplasmic (Potential).
FT TRANSMEM 69 89 2 (Potential).
FT DOMAIN 90 107 Extracellular (Potential).
FT TRANSMEM 108 128 3 (Potential).
FT DOMAIN 129 147 Cytoplasmic (Potential).
FT TRANSMEM 148 168 4 (Potential).
FT DOMAIN 169 202 Extracellular (Potential).
FT TRANSMEM 203 223 5 (Potential).
FT DOMAIN 224 259 Cytoplasmic (Potential).
FT TRANSMEM 260 276 6 (Potential).
FT DOMAIN 277 282 Extracellular (Potential).
FT TRANSMEM 283 302 7 (Potential).
FT DOMAIN 303 345 Cytoplasmic (Potential).
FT DISULFD 105 190 By similarity.
FT CARBOHYD 19 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 345 AA; 38311 MW; DC904127D5B406EC CRC64;

Query Match 40.6%; Score 733; DB 1; Length 345;
Best Local Similarity 42.0%; Pred. No. 5.5e-43;
Matches 141; Conservative 66; Mismatches 127; Indels 2; Gaps 2;

QY 1 MPECHNINISCVNNMNSDVRAISLWVLIITLVGNLIYVISISHFQKLTPTNM 60
DB 10 VLGQCYENVNGSCVKTPTSPGRVLVAVFGGAVLAVFGNLVWISILHFKQLHSPTNF 69
QY 61 LIHSMATVDFLLGCLVMPYSWVRSAGHCWYFGEVFCIKHISTDIMLSASIFHLFSISID 120
DB 70 LIASLACADFWGVGVSPFMSVRSIESCWYFGRSCTFHTCCDVAFCYSSLPHLSFISID 129
QY 121 RYVAVCDPLRYKAKNNIIVICWMIFISWVPAVFAFGMIFELNFKGAEIYKHVHCRG 180
DB 130 RYIAVTDPLVYPTFTYVSVSGICISIWILPLAYSAGAVFYTGVAADGLEEV-SDAVNCVG 188
QY 181 GCSYVFSKISGVLTFMTSFYIPGSIIMLCVYRYIYLIAKEQARLISDANOKLQIGLEMKNG 240
DB 189 GCQVVVNG-NMVLIDFLSFLPTLWMLIGNIFLVARQAKLETGNGKAESSESYKA 247
QY 241 ISQSKERKAVKTLGIWGVFLICWCPFFICTVMDPFLHYIIPTLNDVLIWFGYLNSTFN 300
DB 248 RVARERKRAAKTLGITVAVFMISWLPYSIDSLVDAFMGFTIPAYIYEICWCAVYNSAMN 307
QY 301 PMVYAFYPMFRKALKMLFGKIFOKDSRCKLFLE 336
DB 308 PLIYALFPMFKKAIKIVMSGQVFNSSATMNLFSE 343

Search completed: February 12, 2005, 03:44:40
Job time : 74 secs